AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

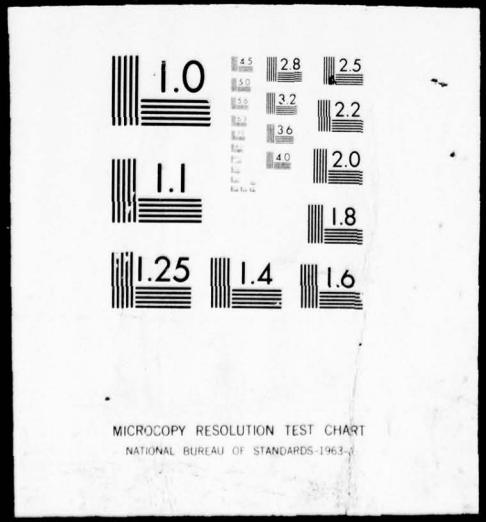
| OF 6
AD A061824

SIFIED

1 OF 6

AD

A051624



1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

Volume I.

Research Reports

Volume I of II

J. Fred O'Brien Jr

Conducted by:
SCHOOL OF ENGINEERING
AUBURN UNIVERSITY
Auburn, Alabama
September, 1977

1977 USAF/ASEE Summer Faculty
RESEARCH PROGRAM

Conducted by
Auburn University
with Assistance from
Ohio State University
and
Other Installations
under
USAF Contract Grant F 44620-75-C-0031

PARTICIPANT'S RESEARCH REPORTS

VOLUME I of II

Submitted to
Air Force Office of Scientific Research
Bolling Air Force Base
Washington, D. C.

by
J. Fred O'Brien, Jr., University Project Director
Associate Director, Engineering Extension Service
Auburn University

September 1977

PREFACE

The USAF-ASEE Summer Faculty Research Program was begun in 1975 with twenty-two members of engineering and science faculties from colleges and universities throughout the country. These professors were assigned to various USAF research laboratories at Wright-Patterson AFB and Eglin AFB for a ten-week period of concentrated research in their selected field and of mutual interest and benefit to the participant (and his university) and the USAF. In 1976, this program was expanded to a total of fifty-three faculty participants assigned to all Air Force Systems Command laboratories. In 1977, the number of participants was expanded to seventy-eight professors.

The basic program objectives are:

(1) To provide scientific and technological benefits to the USAF while enhancing the research interests and capabilities of engineering educators.

(2) To stimulate continuing relations among participating faculty members and their professional peers at the AFSC laboratories.

(3) To form the basis for continuing research of interest to the Air Force at the participant's institution.

(4) To sponsor research in areas of mutual interest to the USAF, the faculty member, and his institution.

The program is conducted under contracts with Auburn University and Ohio State University. The American Society for Engineering Education is co-sponsor of the program.

This document is a compilation of the reports written by participants assigned to laboratories other than Wright-Patterson Air Force Base (Auburn University contract). Mr. J. Fred O'Brien, Jr., Project Director, has exercised certain administrative prerogatives to produce this report.

Similar documentation for the 1975 and 1976 research efforts are on file in the Defense Documentation Center in Washington, D. C. under the following numbers:

1975	Research Reports	ADA031017
1976	Research Reports	ADA033822

The Appendix (in Volume II of this report) contains an index of the 1975 and 1976 research reports.

RESEARCH REPORTS

1977 USAF/ASEE SUMMER FACULTY RESEARCH PROGRAM

RESEARCH REPORTS (Continued)

Report No.	Title	Research Associates
15	Directionally Solidified Eutectic Materials for Electronic, Magnetic, and Optical Applications	Dr. Richard D. Sisson, Jr.
16	Synthesis and Characterization of Novel Nitroaliphatic Materials	Dr. Melvin L. Druelinger
17	Optimal Weapon Assignment in a Tiered Aimpoint System	Dr. Jerren Gould
18	Electrochemical and Luminescence Studies with Special Emphasis on Molten Salt Media	Dr. Csaba P. Keszthelyi
19	Test of Two Hypotheses Concerning the Optical Properties of Cryodeposited Gases	Dr. William B. Newbolt
20	Rocket Exhaust Gas Dynamics	Dr. Peter C. Sukanek
21	Simulation of Turbulent Flow Velocity	Dr. John W. Weatherly, III
VOLUME II		
22	Nitronium Salt Synthesis	Dr. Alton J. Banks
23	Synthesis of Fluorodinitroethylnitramines	Dr. Catherine H. Banks
24	Submicron Particle Size Measurements in Acetylene-Oxygen and Aluminum-Oxygen Flames	Dr. James F. Driscoll
25	Solid Teflon Propulsion Contamination Investigation	Dr. James K. Hartman
26	Infrared, Raman, Mass Spectroscopy and Gas Chromatography of Selected Fluorodinitro Plasticizers	Dr. Steven Riethmiller
27	Rotational and Mean Flow Effects on Motor Stability	Dr. Wm. K. Van Moorhem
28	Investigation of Raman Microwave Emission from Plasmas	Dr. Tse-Wen Hsu
29	Automatic Speech Processing-Speech Enhancement	Dr. Russell J. Niederjohn
30	Design of System Diagnostic and Fault Isolation Procedures	Dr. Theodore J. Sheskin

RESEARCH REPORTS (Continued)

PARTICIPANT LABORATORY ASSIGNMENT

1977 USAF/ASEE SUMMER FACULTY RESEARCH PROGRAM

AFCEC — AIR FORCE CIVIL ENGINEERING CENTER (TYNDALL AIR FORCE BASE)
1. Harry H. Caldwell - University of Idaho
2. Braja M. Das - South Dakota State University
3. John A. Stevens - Youngstown State University

AFCEEDO — AIR FORCE CIVIL AND ENVIRONMENTAL ENGINEERING DEVELOPMENT OFFICE (TYNDALL AIR FORCE BASE)
1. Charles Springer - University of Arkansas
2. Alvin M. Strauss - University of Cincinnati
3. Paul H. Zedler - San Diego State University

AFATL — AIR FORCE ARMAMENT LABORATORY (EGLIN AIR FORCE BASE)
1. Jan F. Andrus - University of New Orleans
2. Maurice J. Bouvier - Louisiana State University
3. Robert W. Courter - Louisiana State University
4. Carl Grant Looney - University of Toledo

DET — ROME AIR DEVELOPMENT LABORATORY (HANSCOM AIR FORCE BASE)
1. Vaughn P. Cable - California State University
2. Richard Dean Sisson, Jr. - Worcester Polytechnic Institute

ESD — ELECTRONICS SYSTEMS DIVISION (HANSCOM AIR FORCE BASE)
1. Marvin D. Drake - Florida Institute of Technology

AFGL — AIR FORCE GEOPHYSICS LABORATORY (HANSCOM AIR FORCE BASE)
1. Thomas E. Morgan - State University of New York
2. James R. Sharber - Florida Institute of Technology

FJSRL — FRANK J. SEILER RESEARCH LABORATORY (AIR FORCE ACADEMY)
1. Melvin L. Druelinger - Indiana State University
2. Jerren Gould - Claremont Graduate School
3. Csaba Peter Keszthelyi - Louisiana State University

AEDC — ARNOLD ENGINEERING DEVELOPMENT CENTER (ARNOLD AIR FORCE BASE)
1. William B. Newbolt - Washington & Lee University
2. Peter Sukanek - Clarkson College of Technology
3. John Weatherly III - Louisiana State University

AFRPL — AIR FORCE ROCKET PROPULSION LABORATORY (EDWARDS AIR FORCE BASE)
1. Alton J. Banks - Texas A & M University
2. Catherine H. Banks - Texas A & M University
3. James F. Driscoll - University of Michigan
4. J. Keith Hartman - Canisius College
5. Steven Riethmiller - Virginia Military Institute
6. William K. Van Moorhem - University of Utah

PARTICIPANT LABORATORY ASSIGNMENT Continued

RADC — ROME AIR DEVELOPMENT CENTER (GRIFFISS AIR FORCE BASE)

1. Tse-Wen Hsu - Villanova University
2. Russell J. Niederjohn - Marquett University
3. Theodore J. Sheskin - Cleveland State University
4. Robert G. Van Meter - State University College
5. Ralph M. Weischedel - University of Delaware

SAM — SCHOOL OF AEROSPACE MEDICINE (BROOKS AIR FORCE BASE)

1. Thomas J. Jones - Trinity University
2. Demetrius D. Venable - St. Paul's College
3. James M. Verlander - Texas A & M University

AFHRL-FT — AIR FORCE HUMAN RESOURCES LABORATORY (WILLIAMS AIR FORCE BASE)

1. Herbert A. Colle - Wright State University

AFHRL-FT — AIR FORCE HUMAN RESOURCES LABORATORY (LUKE AIR FORCE BASE)

1. John A. Seevers - Washington State University

AFWL — AIR FORCE WEAPONS LABORATORY (KIRTLAND AIR FORCE BASE)

1. Richard A. Anderson - University of Missouri-Rolla
2. Martin D. Bradshaw - University of New Mexico
3. Wayne A. Charlie - Colorado State University
4. Richard L. Redington - Texas Tech University
5. Harold D. Southward - University of New Mexico
6. Stephen G. Wells - University of New Mexico

LIST OF PARTICIPANTS

1977 USAF/ASEE SUMMER FACULTY RESEARCH PROGRAM

AFCEC/CEEDO (Tyndall)	DEGREE, SPECIALTY & LAB ASSIGNMENT
Dr. Harry H. Caldwell Department of Geography University of Idaho Moscow, ID 83843	Degree: Ph.D., Geography Specialty: Livability Meas Systems Assigned: AFCEC Colleague: Mr. Charles Lewis
Dr. Braja Mohan Das Associate Professor Civil Engineering Department South Dakota State University Brookings, SD 57006	Degree: Ph.D., Soil Mech & Foundations Specialty: Soils and Foundations Assigned: AFCEC Colleague: Major James I. Clark
Dr. Charles Springer Professor, Chemical Engineering Dept 332 Engineering Building University of Arkansas Fayetteville, AR 72701	Degree: Ph.D., Chemical Engineering Specialty: Mass Transfer Assigned: DET I (CEEDO) HdQtrs. ADTC Colleague: Major Peter Daley
Dr. John Augustus Stevens Associate Professor Chemical Engineer & Material Science Dept The Rayen School of Engineering Youngstown State University Youngstown, OH 44555	Degree: Ph.D., Physical Chemistry Specialty: Corrosion and Materials Assigned: AFCEC Colleague: Mr. Harold Stevens Mr. Tom Lewicki
Dr. Alvin M. Strauss Engineering Science Department University of Cincinnati Cincinnati, OH 45221	Degree: Ph.D., Theor. & Appld. Mech. Specialty: Mechanics, Applied Math Assigned: DET I (CEEDO) HDQTRS ADTC Colleague: Capt. Monte H. Callen
Dr. Paul H. Zedler Department of Biology San Diego State University San Diego, CA 92182	Degree: Ph.D., Botony Specialty: Plant Ecology Assigned: DET I (CEEDO) HDQTRS. ADTC Colleague: Major R. C. Wooten
AFATL (Eglin)	
Dr. Jan Frederick Andrus Associate Professor, Dept of Math University of New Orleans New Orleans, LA 70122	Degree: Ph.D., Math Specialty: Numerical Analysis Assigned: AFATL Colleague: Mr. Mike Caluda
Dr. Maurice J. Bouvier Electrical Engineering Department Louisiana State University Baton Rouge, LA 70803	Degree: Ph.D., Elec Engineering Specialty: Communications Systems Assigned: AFATL Colleague: Major Dave DePriest

AFATL (Eglin) Continued

Dr. Robert W. Courter
Associate Professor, Mech Engineering
Louisiana State University
Baton Rouge, LA 70803

Degree: Ph.D., Psychology
Specialty: Aerodynamics, Flight Dyn
Assigned: AFATL
Colleague: Mr. Ken Cobb

Dr. Carl Grant Looney
Assistant Professor
Mathematics Department
University of Toledo
Toledo, OH 43603

Degree: Ph.D., Mathematics
Specialty: Computational Mathematics
Assigned: AFATL
Colleague: Mr. Aaron Brinson

DET/ESD/AFGL (Hanscom)

Dr. Vaughn Paul Cable
Assistant Professor
Dept of Electrical & Electronics Engineering
California State University

Degree: Ph.D., Electromagnetic Theory
Specialty: Electromagnetic Theory
Assigned: DET
Colleague: Dr. R. J. Mailloux

Dr. Marvin D. Drake
Assistant Professor, Dept of Elec Engr.
Florida Institute of Technology
P. O. Box 1150
Melbourne, FL 32901

Degree: Ph.D., Electrical Engineering
Specialty: Electro-optic, Mat'ls & divices
Assigned: ESD(DP)
Colleague: Dr. Donald B. Brick

Dr. Thomas E. Morgan
Acting Chairman
Department of Earth Sciences
Oswego, NY 13126

Degree: Ph.D., Astrophysics
Specialty: Solar & Stellar Atmos
Assigned: AFGL
Colleague: Mr. Charles Dubs

Dr. James Randall Sharber
Physics and Space Science Deparmtnet
Florida Institute of Technology
Melbourne, FL 32901

Degree: Ph.D., Physics
Specialty: Near-earth space physics
Assigned: AFGL
Colleague: Dr. James A. Whalen

Dr. Richard Dean Sisson, Jr.
Assistant Professor
Mechanical Engineering Materials Division
Worcester Polytechnic Institute
Worcester, MA 01609

Degree: Ph.D., Materials Engineering
Specialty: Materials Engineering
Assigned: DET
Colleague: Mr. Alton Armington

FJSRL (AF Academy)

Dr. Melvin Leroy Druelinger
Associate Professor
Chemistry Department
Indiana State University
Terre Haute, In. 47809

Degree: Ph.D., Organic Chemistry
Specialty: Organic Chemistry
Assigned: FJSRL
Colleague: Capt. Scott Shackleford

FJSRL (AF Academy) Continued

Dr. Jerren Gould
Mathematics
Claremont Graduate School
Claremont, CA 91711

Degree: Ph.D. Statistics
Specialty: Statistical Theory, O'n's Res.
Assigned: FJSRL
Colleague: Capt. S. J. Monaco

Dr. Csaba Peter Kexzthelyi
Assistant Professor
Chemistry Department
Louisiana State University
Baton Rouge, LA 70803

Degree: Ph.D., Electrochemistry
Specialty: Non-Aqueous Systems
Assigned: FJSRL
Colleague: Capt. C. L. Hussey
Capt. L. Davis

AEDC (Arnold)

Dr. William Barlow Newbolt
Professor of Physics
Washington and Lee University
Lexington, VA 24450

Degree: Ph.D., Physics
Specialty: Nuclear Spectroscopy
Assigned: AEDC
Colleague: Mr. J. A. Roux

Dr. Peter Sukanek
Clarkson College of Technology
Potsdam, NY 13676

Degree: Ph.D., Chemical Engineering
Specialty: Fluid Mechanics
Assigned: AEDC
Colleague: Mr. R. Rholes

Dr. John Wilson Weatherly III
Associate Professor
Mechanical & Ind. Engineering Dept
Louisiana State University
Baton Rouge, LA 70803

Degree: Ph.D., Mechanical Engineering
Specialty: Thermal Sciences
Assigned: AEDC
Colleague: Dr. Don Barnett

AFRPL (Edwards)

Dr. Alton Joseph Banks
Department of Chemistry
Texas A & M
Texas Avenue
College Station, TX 77843

Degree: Ph.D., Chemistry
Specialty: Inorganic Chemistry
Assigned: AFRPL
Colleague: Dr. Claude Merrill

Dr. Catherine Hamrick Banks
Department of Chemistry
Texas A & M
Texas Avenue
College Station, TX 77843

Degree: Ph.D., Organic Chemistry
Specialty: Organo-Sulfur Compounds
Assigned: AFRPL
Colleague: Dr. Claud Merrill

AFRPL (Edwards) Continued

Dr. James Francis Driscoll
Assistant Professor
Department of Aerospace Engineering
209 Aerospace Engineering Building
The University of Michigan
Ann Arbor, MI 48105

Degree: Ph.D., Aerospace & Mech Scienc
Specialty: Laser Meas tech. Gasdynamics
Assigned: AFRPL
Colleague: Dr. David Mann

Dr. James Keith Hartman
Physics Department
Canisius College
Buffalo, NY 14208

Degree: Ph.D., Applied Physics
Specialty: Modeling of Physical Systems
Assigned: AFRPL
Colleague: Lt. Sharon Pruitt

Dr. Steven Riethmiller
Associate Professor
V.M.I. Chemistry Department
Lexington, VA 24450

Degree: Ph.D., Chemistry
Specialty: Molecular Spectroscopy
Assigned: AFRPL
Colleague: Mr. Louis Dee

Dr. William K. Van Moorhem
Mechanical Engineering Department
University of Utah
Salt Lake City, UT 84112

Degree: Ph.D., Aerospace Engineering
Specialty: Fluid Mechanics, Acoustics
Assigned: AFRPL
Colleague: Mr. Jay Levine

RADC (Griffiss)

Dr. Tse-Wen Hsu
Electrical Engineering Department
Assistant Professor
Villanova University
Villanova, Pennsylvania 19085

Degree: Ph.D., Microwave Electronics
Specialty: Microwave Electronics
Assigned: RADC
Colleague: Mr. R. H. Chilton

Dr. Russell James Niederjohn
Associate Professor
Electrical Engineering Department
Marquett University
1515 W. Wisconsin Avenue
Milwaukee, Wisconsin 53233

Degree: Ph.D., Electrical Engineering
Specialty: Speech Processing
Assigned: RADC
Colleague: Capt Robert Curtis

Dr. Theodore Jerome Sheskin
Assistant Professor of Industrial Engineering
Cleveland State University
Cleveland, OH 44115

Degree: Ph.D., Industrial Engineering
Specialty: Operations Research
Assigned: RADC
Colleague: Mr. Jerry Klion

RADC (Griffiss) Continued

Dr. Robert Guy Van Meter
Department of Mathematical Sciences
State University College
Oneonta, NY 13861

Degree: Ph.D., Mathematics (Algebra)
Specialty: Finite fields
Assigned: RADC
Colleague: Capt Kenneth E. Wilson

Dr. Ralph M. Weischedel
Assistant Professor
Stat & Computer Science
University of Delaware
Newark, DE 19711

Degree: Ph.D., Computer Science
Specialty: Artificial Intelligence
Assigned: RADC
Colleague: Mr. William E. Rzepka

SAM (Brooks)

Dr. Thomas Jefferson Jones
Associate Professor
Engineering Department
Trinity University
San Antonio, TX 78284

Degree: Sc.D., Electrical Engineering
Specialty: Electronic Control Systems
Assigned: SAM
Collegue: Mr. Kent Gillingham

Dr. Demetrius Dante Venable
Assistant Professor Of Physics
Saint Paul's College
Lawrenceville, VA 23868

Degree: Ph.D., Physics
Specialty: Physics
Assigned: SAM
Colleague: Dr. John Toboada

Dr. James Milton Verlander, Lecturer
Department of Biology
Texas A & M University
College Station, TX 77843

Degree: Ph.D., Physiology
Specialty: Elect Activity of Brain
Assigned: SAM
Colleague: Major Gibbons

AFHRL-FT (Williams)

Dr. Hergert Anthony Colle
Department of Psychology
Wright State University
Dayton, OH 45431

Degree: Ph.D., Psychology
Specialty: Psych; Human Performance
Assigned: AFHRL-FT
Colleague: Mr. James Smith

AFHRL-FT (Luke)

Dr. John Altman Seevers
Department of Mechanical Engineering
Washington State University
Pullman, Washington 99163

Degree: Ph.D., Engineering
Specialty: Dynamic Systems Analysis
Assigned: AFHRL-FT (Luke)
Colleague: Mr. Bill Hopkins

AFWL (Kirtland)

Dr. Richard Alan Anderson Professor of Physics Physics Department University of MO-Rolla Rolla, MO 65401	Degree: Specialty: Assigned: Colleague:	Ph.D., Physics Physics and Spectroscopy AFWL Mr. Steve Davis
Dr. Martin D. Bradshaw EECS Department University of New Mexico Albuquerque, NM 87131	Degree: Specialty: Assigned: Colleague:	Ph.D., Elec Engineering Electromagnetic Fields AFWL Dr. J. Phillip Castillo
Dr. Wayne Alexander Charlie Department of Civil Engineering Colorado State University Fort Collins, CO 80523	Degree: Specialty: Assigned: Colleague:	Ph.D., Civil Engineering Geotechnical Engineering AFWL
Dr. Richard Lee Redington Professor of Chemistry Texas Tech University Lubbock, TX 79409	Degree: Specialty: Assigned: Colleague:	Ph.D., Physical Chemistry Molecular Structure AFWL Mr. Leroy Wilson
Dr. Harold D. Southward Professor, Electrical Engineering Dept University of New Mexico Albuquerque, NM 87131	Degree: Specialty: Assigned: Colleague:	Ph.D., Physics Solid State Physics & Devices AFWL Lt Col Peter D. Tannen
Dr. Stephen Gene Wells Department of Geology University of New Mexico Albuquerque, NM 87131	Degree: Specialty: Assigned: Colleague:	Ph.D., Geology Geomorphology/Surficial Pro. AFWL Capt Alber R. Schenker, Jr.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

LIVABILITY AS APPLIED TO AIR FORCE BASES

Prepared by:	Harry H. Caldwell, Ph.D
Academic Rank:	Professor
Department and University:	Department of Geography University of Idaho

Assignment:	
(Air Force Base)	Tyndall AFB
(Laboratory)	Civil Engineering Center
(Division)	Environmental Planning
(Branch)	Community Development
USAF Research Colleague:	Charles F. Lewis, M.U.P.
Date:	August 12, 1977
Contract No.	F44620-75-C-0031

LIVABILITY AS APPLIED TO AIR FORCE BASES

by

Harry H. Caldwell

ABSTRACT

The Air Base has numerous parallels to the isolated small company town and surrounding communities proximate resort, construction or boom towns. Small towns permit little privacy, or extremes, have a strong sense of community,and show great interest and concern about morality, behavior, and local competitive sports.

There are significant differences between the civilian Quality of Life (QOL) concept and livability as applied to the military as some elements do not fit and others are provided. The Air Force is committed to improve livability and has an impressive performance record though wants are insatiable.

Bases have many functions that influence livability. BCE has a special role in dealing with built environments. The base commander, social action, chaplain, advisory councils, legal groups and the Hospital handle other facets - but without common goals. Livability perceptions vary between individuals and with a person over time. They are comparative to prior experiences and it is easier to upgrade than downgrade elements of comfort, convenience and privacy. With more off base housing, job attitudes, commuting and regular work hours, there has been a decline in the sense of community.

Detailed livability indices for base housing and working environments vary with housing types and work units and create unwieldy numbers of entities. Livability should be regarded as a people sensitive dimension of comprehensive planning and decision makers should weigh livability considerations against readiness, mission effectiveness and productivity.

ACKNOWLEDGMENTS

After 25 years of study devoted to livability, environmental and planning problems of small towns and non-metropolitan regions, it was gratifying to spend a concentrated 10 week period on a first hand examination of Air Force Base Livability Problems. For this opportunity I am indebted to the USAF-ASEE program, the staff of J. Fred O'Brien, Program Director, photographic, secretarial, library and graphics personnel at Tyndall and my cooperative research colleague, Charles F.Lewis,as well as others in Community Planning and in the Directorate of Environmental Planning at the Air Force Civil Engineering Center.

Special thanks are due Lester Henriksen at USAF-PREV for encouraging me to apply and for arranging my briefings to division chiefs, and to Generals Thompson and Gilbert at Hq USAF/PRE. The credit for deciphering my handwriting and typing the manuscript is due Myrdel "Mike" Wilson, a most cooperative and accommodating secretary at AFCEC.

LIST OF TABLES

INTRODUCTION

The U.S. Air Force has a long history of concerns for livability. Numerous programs, benefits and facilities have been provided. Most of these are service wide or even multi-service. At the base level many different units address themselves to livability questions including the base commander, his various advisory councils, unit commanders,first sergeants andactually all supervisors are asked to be sensitive to the individual needs and aspirations of people in their unit - both military and civilian.

Historically, the cliche that evolved to accommodate complaints of inequity or insensitivity was to "tell it to the chaplain," a route still popular as it provides for privileged information and confidentiality. The other traditional outlet was the Inspector General. There are responses possible through Human Relations, Social Action and at the Hospital for mental as well as physical health concerns.

Base Civil Engineering has a major livability role in the design and location of structures and facilities including commissary, housing, work areas, BX, streets, traffic flow, laundry, heating, cooling, maintenance, solid wastes, sewer and mortuary affairs. Educational services, billeting, recreation and library facilities, programs and opportunities enrich or detract from the base livability.

Regulations impacting livability range from those of the unit commander up to the Office of Secretary of Defense and even Congress which establishes control by funding selectively and with stipulations or conditions. Administrative interpretations and attitudes of these regulations are critical.

The Air Force provides well for its military personnel and dependents with medical care including births, child care, youth centers, housing facilities and allowances, annual and sick leave, commissary system, educational opportunities and benefits, counseling,early retirement, unique retirement privileges, widow and dependent homes and villages and even burial rights.

However, the priority of livability in the heirarchy of Air Force goals is unclear. The primary goals are readiness and mission effectiveness which translates to the most modern sophisticated equipment in terms of planes, missiles, drones, armor andrapid global response in conjunction with the other services. On bases, this places the highest priority on the runway and a special role for pilots and flight crews.

Secondary goals include livability along with cost effectiveness, productivity, energy and resource conservation, environmental protection, flexibility, aesthetics, and compatible community relations. There are other specific goals such as increasing the number of women in the service and in upgrading of minorities commensurate with their ability and performance.

It is evident to this author that the secondary goals have not been prioritized and that many supervisors are not certain which ones should receive lip service, token support, or full compliance. Despite proclamations, posters,notices, and memoranda, practices on specific bases do not transmit the same priority message. This results in reduced credibility - internally and externally, a reduced sense of community or pride in the service, more self centered behavior, and less willingness to make personal or family sacrifices to the mission. In recent years there were improvement in the pay structure and financial incentives provided for reenlistment as the military made the transition from a draft to volunteer service. Though the intent was to insure adequate numbers it also enhanced livability though not necessarily greater productivity, readiness or mission effectiveness.

At the start of this project it was assumed possible to quantify and establish separate objective livability index values for different activities on a base. Housing and work areas were selected for study, using the local base as a test case. Literature searches provided rich information about the conditions that enhance or detract from housing and work livability and ways to determine satisfaction and dissatisfaction levels.

Models that apply to a single civilian company were not transferable to air force bases because working conditions vary greatly in different support areas, between the mission oriented and support groups and between civilians and military personnel. This is a training base and there are large influxes of new groups in for short stays which further prevents meaningful quantification. The end result of this search was an analysis of the components of base work livability but a rejection of the concept of a separate work livability index.

The housing study raised different questions as it was immediately restricted to bachelor and family base housing. A tri-service occupant study of married base housing made in 1972 provided valuable but dated information. Some of this was updated by a 1975 Air Force Housing Study which dealt with some livability considerations in the original construction and in subsequent rehabilitation.

Some livability indexing came from the 1972 survey for each major category of housing such as the Wherry and Capehart units

which are common to most bases. Other houses built by special appropriated funds since 1963 are base specific and too variable in design and construction except for local indexing. The process gets more complicated as officer houses tend to be larger and to have more closet and storage space than the housing of enlisted personnel. Though there are some 150,000 units of married base housing, there are at least an equal if not greater number of married military personnel who live off the bases and they would not be covered with a housing livability index. Thus, in family housing as in the work areas, it was deemed more meaningful to establish guidelines for identifying livability concerns than to construct separate quantified evaluations.

Bachelor housing has its own sets of variables depending whether the quarters are old or new, 1, 2 or 3 stories, spacious or cramped, and whether the first sergeant is demanding or permissive and if the quarters are for permanent staff, temporary duty groups, men or women. New bachelor officer quarters along with the 2 or 3 story enlisted dormitories tend to favor motel type construction, with external entrances and little emphasis on communal rooms. These designs favor high privacy values, and opportunities to individualize rooms but they work against the unifying sense of a common community. A tentative bachelor housing livability index system was developed and is presented.

BASE LIVABILITY AND THE SMALL TOWN MODEL

This base livability study relates mainly to military personnel, though civilians are an important segment of every base. Military bases do not exist in a vacuum and are closely associated with surrounding communities that provide housing for all the civilian employees, and for many military and their dependents. The adjacent areas also provide many leisure time pursuits, services, commercial establishments, educational and cultural opportunities that are shared mutually.

Military bases show numerous parallels to isolated small towns, expecially the company town where one concern has built the town, the houses, stores, schools, factory and basic utilities. The company makes and enforces the rules. The first priority is profit maximization which is the counterpart of 'readiness' and mission effectiveness. Housing modernization ranks well below productivity, and housing restrictions, despite low rental charges, make home ownership off the property desirable for many in both situations. Small towns adjacent to air bases tend to undergo growth and structural change that parallels resort or boom towns. Gradually the adjacent town tends to move certain functions closer to air bases to shorter 'desire lines' for working commuters, military and civilian, in

regard to housing and retirees with BX, recreation and commissary privileges. Here, the law of intervening opportunities becomes operative as the corridor routes from town to base host a plethora of commercial outlets and billboards.

The following characteristics of small towns may or may not apply to Air Force bases:

They have a sense of community but little privacy or individuality. The populace assembles for watermelon feeds, pancake breakfasts, volunteer fire fighters and interdenominational churches. Often accused of prejudice and provincialism, these same shortcomings are found equally in larger cities. Small towns provide more fusion than friction despite wide divergences of religion, values and politics. They are more concerned with morality than with the transmission of wisdom, truth or knowledge. The latter three may be regarded as less relevant than the high school football and basketball teams that provide community drive, motivation and team spirit.

The small town offers a peaceful and calm environment, but many young people want something other than peace and calm. Small towns and even the anti-metropolitan attitude are part of the American heritage. Many small towns resist growth and are fearful of change and the influx of new, strange people. Change is accepted slowly and only certain individuals are accepted as innovators as part of their status role. In small company towns, extremes are not tolerated - only moderation. Extremes might threaten the power structure and produce job loss repercussions. With the combination of corporate control,paternalism, and a security orientation, company towns do not tend to mature and evolve as strong, sound, responsible communities. Instead, they tend to stagnate in an adolescent role with an undercurrent of rebellious antagonism. In the company town (base parallel) one gives up some independence and responsibility for security.

Some small towns have traded their traditional values for prosperity, tourism and new factories. They run the full gamut of ski towns, construction towns (boom and bust places), irrigation, dam, historic and resort towns.

Some start out with great plans and ideals but soon reflect stereotyped house types, with lawns and flowers transplanted to the new locations. People who were attracted by the green hills of an area immediately try to reduce the amount of green. This is similar to some older base housing areas where the trees were first leveled to build homes and then funds were sought to provide shrubbery.

Resorts, boom towns and some towns adjacent to military bases are unstable, plastic places with changing people,jobs and

activities. They are characterized by much manipulation, political land deals and cultural poverty, as high investment returns is the primary goal.

There are a number of dynamic considerations in this model that result in more people choosing to live off the base, to regard the base as the job location, a place where friends reside, to regard the commissary as a type of discount supermarket and only reluctantly return to the base for evening social activities at the appropriate club if the commuting time and distance is not excessive.

Though some small towns continue to thrive in the United States, the company town has virtually disappeared except for those in isolated locations. Further study on the parallels to air force bases might provide additional behavior and trend insights.

QUALITY OF LIFE AND BASE LIVABILITY

In the wake of environmental and social alarms in the 1960's the phrase 'quality of life' surfaced widely with diverse meanings. From 1968 to 1972 several studies came from individuals, committees, conferences, EPA Fellows and consulting firms that sought to define, analyze or quantify quality of life (QOL), and even rank cities on diverse sets of indicator criteria. While each approach had a different emphasis reflecting the focus of the group or their client, there was significant overlap.

The Environmental Studies Division of the Environmental Protection Agency desired a quantitative Quality of Life Index to evaluate socio-cultural elements of Environmental Impact Statements. An elite group of 150 specialists was assembled to identify the components and weightfactors in terms of their importance.[1] Though they did not agree upon a quantitative measuring device, their collective weighting of 47 factors did produce a quality of life ranking system (Table I) and the top 11 items are listed. When the 47 factors are grouped into larger groupings with similar weights, the main categories were Economic - 32%, Political/Social - 36%, and Environmental - 31%. Environmental concerns were viewed as a group but component units did not rank high on this list. A notable ommission were the psychological stress factors behind urban law and order concerns.

Several of the highly rated QOL factors either do not apply to the military (as democratic processes) or else they have been institutionalized (providing health care). Provisions have been made for public participation in both the advisory role (advisory councils) and in adversary situations (social action, Inspector General).

T A B L E I

QUALITY OF LIFE
1971 Ranked Factors

1. Democratic Processes
2. Public Participation
3. Health
4. Choices in Life
5. Housing
6. Economic Security
7. Education
8. Land Use
9. Living Costs
10. Economic Opportunity
11. Ecosystem

*EPA Conference 1971

In 1973 another study on Quality of Life[2] for EPA (Table II) established and outlined some different factors as well as reporting on the findings in six other QOL studies. Though these items were less conceptual than the earlier EPA study, it went beyond perception to identify valid interacting elements. Many of these factors are not prime concerns at military installations - unless there was a local problem. Several base livability concerns were not specifically addressed in these broad spectrum studies.

T A B L E II

MAJOR QUALITY OF LIFE FACTORS - EPA-1973

I. Economic Environment
   - Income
   - Income Distribution
   - Economic Security
   - Work Satisfaction

II. Social Sector
   - Family
   - Community
   - Social Stability
   - Physical Security
   - Culture
   - Recreation

III. Political Environment
- Electoral Participation
- Non-electoral Participation
- Government Responsibility
- Civil Liberties
- Informed Constituency

IV. Health
- Physical
- Mental
- Nourishment

V. Physical Environment
- Housing
- Transportation
- Public Services
- Material Quality (goods and services)

VI. Natural Environment
- Air and Water Quality
- Radiation
- Toxicity
- Solid Wastes
- Noise

From studies, reports, regulations, observations and interviews, an unranked list of Base Livability Factors have been tabulated. Some of these factors interconnect with the region and adjacent communities, others reflect personal values, and family concerns, spatial components, short and long term elements.

T A B L E III

BASE LIVABILITY-1977
Unranked Factors

*	Housing
	Economic Security
**	Leisure Time Facilities
*	Educational Opportunities
**	Physical/Environmental
	Advancement Opportunities
	Administrative Practices
	Rank/Status Consciousness
	Personal Acceptance
	Privacy
	Relative Equity
	Meaningful Work

* In Commuting Area
** Regional Context

The EPA QOL lists contain broad societal or conceptual concerns but the Base Livability list focuses on the applied living-working-leisure time environments of military personnel. If AF civilian employees were considered, there would be concern with rights and privileges on the base, commuting time, and military-civilian relations. Dependents would add items close to their daily lives.

Much of the public alarm about Quality of Life subsided with the creation of such institutional responses as the creation of EPA, the passage of the other environmental and social laws, including the Freedom of Information Act, laws requiring open meetings, and public participation at various levels. Based on prior collective behavior patterns in America, institutionalization ameliorates concerns.

Livability as perceived at military bases is different in focus and priorities than the broader Quality of Life concerns.

Though traditional tours of duty last 2-3 years, there is much additional movement tied to special schooling and temporary duty that creates an atmospher of nomadic transients. This military mobility is unlike civilian mobility in that it is involuntary, frequently requires family separation and generally is accompanied by a mobile environment, creating 'little America ghettos' overseas intact with Cub Scouts, Barbeques, and a Santa Claus at the BX.

"Many American families living overseas in American style housing and ghetto communities pass up priceless opportunities for involvement in another culture and for enriching experiences." "To live in a charming tatami-matted, paper-walled, non-heated house among neighbors who speak only Japanese can be a delightful experience if you choose it; to have it forced on you is quite another matter. As long as military mobility remains involuntary, American type communities for those serving overseas will be a necessity in order to maintain an acceptable level of morale."[3]

This identifies another uniquely military dimension of livability - where the social system is composed of the institutional structures, regulations, administrative practices, and technological capacities.

LIVABILITY PROBLEMS

The casual non-military observer might come to regard military livability as a never ending list of wants and benefits during service and after retirement wants are insatiable. Even within the organization there is some envy generated between ranks and between civilian employees and the military.

There are military lobbyists and supporters trying to obtain legislation or rulings to enlarge benefits or to lower the eligibility barriers to existing benefits. This constitutes one area of livability with a backlash potential.

There is another problem in dealing with livability as though it were a complete goal in itself instead of one important dimension in the Comprehensive Planning process that should be matched to productivity, readiness, mission and cost effectiveness.

Recommendation - that livability considerations be tested to insure that they will not reduce productivity, mission effectiveness and response capability and that they be cost effective. The cost effective measure must include the psychological, social, aethestic benefits and costs as well as traditional economic measures.

There are still many shortfall areas in livability that can be resolved with minimal cost provided that planning and decision making is done in a framework that people and their feelings matter.

MILITARY LIVABILITY VALUES AND FRAMEWORKS

In their award winning paper, Henriksen and Vest, defined livability as "the total of all the physical, administrative, regulatory and socio-environmental events or circumstances that the individual encounters in the course of his or her daily life."[4]

The main omission is in the opening phrase "the perception of the impact on the individual, family and friends, of etc." A livability index represents the congruence or gap between some ideal set of conditions and the immediate conditions that they perceive in their place and situation.

The difficulty in indexing is that the evaluation comes from a frame of reference established through prior experience in previous environments. If the experience at basic training is spartan and demanding, it establishes the first reference frame for military livability. It is easy to accept upgraded quarters and facilities but difficult to accept more austere conditions except at remote posts or for short periods. This leads to a recommended progression of upgraded facilities during a tour of duty probably geared to the number of years in service rather than by rank or family size. If this is not feasible, then basic training posts and first assignments should provide comfortable but more sparse accommodations than the norm for the service.

In seeking the common denominators used in livability perception frameworks, the items in Table IV seem most significant.

T A B L E IV

PERCEPTION FRAMEWORK INGREDIENTS

Rumor and Hearsay from Peer Groups

Impacts on Family Life

Sense of Acceptance (self and dependents)

Perceived and Judged Physical Environment

Comfort and Convenience

Work and Housing Satisfactions

Sense of Personal Growth

Several Types of Privacy

Reliability and Trust in Supervisors and Institution

Institutional Loyalty or Sense of Community

Ingredients with special roles are described in additional detail.

Comfort and Convenience include physiological, psychological and social needs. These are escalating values that reflect peer group pressures, technology, promotional advertising, and middle and upper class civilian comparisons, all set in the world of rising comfort expectations. At isolated posts and in emergency or temporary conditions, individuals seem willing to endure lower comforts. Selected groups, individuals, and periodically total societies will voluntarily accept less comfort and convenience when highly motivated by personal and social commitment to social, religious, political, economic or even ecologic goals. Over long time periods, defections surface and the movement toward comforts and conveniences is reestablished. Comfort and convenience can be indexed from satisfaction/dissatisfaction levels using comparative data as was done in the 1972 occupant survey of military housing.

Another ingredient is a Sense of Community, an extension of the common Air Force goal, a type of extended family, the sense of a career, or a calling in which the military and their families feel a strong sense of obligation, commitment and acceptance. During World War II and in the Korean conflict, the sense of community was quite visible and still exists amongst senior personnel that have known the feeling that develops under combat conditions or in the sharing of a meaningful mission role in

remote duty locations. Some structural and administrative changes designed to improve aspects of livability have proven damaging to the sense of community.

Increased salaries, bonuses for reenlistment, and the increased popularity of off base housing for married and bachelor military, increased privacy and individualization,along with the near demise of basic training, and the end of open bay barracks, result in a decreased sense of community. This change accelerates the move to regard the service more along job or occupational lines.

A sense of equitable treatment, acceptance by one's peers, neighbors and the community and opportunities for individual differences and variety are additional important values.

Another frequently cited though not fully appreciated concept are the numerous elements involved in privacy. Privacy is an interpersonal boundary control process. It involves selective control, the opening and closing of the self or groups to and from others. It may involve the physical self or emotions or feelings. It may be visual, accoustical emotional or informational. Most common is the notion of Solitude - to be alone and free from observation by others. Equally important is Intimacy as in the separation of a small group (husband-wife) from others. Anonymity as being "lost in a crowd" and not being recognized is another form of privacy as is Reserve - in the form of a psychological barrier against unwanted intrusion.[5] The use of headphones with stereo sets, TVs, etc., represents accoustical privacy.

The functions of privacy vary with the individual and situation. It produces personal autonomy which relates to opportunities for concentration, self-independence, self reliance and self identity. It may also provide an emotional release - or relaxation from social roles, to be "off stage" and to deviate from rules, customs, traditions and norms in a protected fashion. It provides an atmosphere for self evaluation, integration of experiences, contemplation and the opportunity to plan and assess future actions. In designing work environments spaces for limited and protected communication are needed to share confidences with individuals or small groups (i.e., supervisor discussions).

Privacy is a continually changing process with needs that vary hourly, daily, seasonally and during one's life and which reflect changing sensitivity dependent upon personal histories and cultural backgrounds.

As one seeks pragmatic privacy approaches to livability, the interface between privacy and a sense of community or acceptance emerges. It is derived from comparing achieved privacy with desired privacy, the latter being the ideal state perceived

by individuals or groups concerning social interaction. When achieved privacy is less than desired privacy, the result is intrusion or invasion of privacy. When achieved privacy is greater than desired privacy the results may be called foredom, loneliness, social isolation or alienation. There are parallels in desired and achieved livability.

The techniques people use to implement their desired levels of privacy are verbal, paraverbal, behavior, nonverbal body language, environmental behaviors and cultural norms - all functioning as an integrated system. A head nod, smile and verbal praise may be used in conjunction and complement each other to indicate strong agreement. The open door, drapes, raised window shade, extra office chairs and open desk arrangement may say welcome and that privacy barriers are low. However, that day or week the individual may feel the need for considerable privacy this confuses the outsider because the different privacy mechanisms are not giving the same message.

The inability to maintain interpersonal privacy boundaries can produce certain costs - expressed as physical work-effort-stress, psychological energy expenditures stress, tension and anxiety.

Continued invasions of privacy may endanger personal autonomy, self-respect and dignity. In the process it reduces a person's functional efficiency.

PERCEPTION DYNAMICS

Individuals perceive livability changes differently with maturity, sex, family status, critical ages in children, experience and with changing interests in group activities compared to the importance of privacy.

Interest in personal acceptance and recognition by peer groups to a first termer is higher than with senior NCOs. Livability for the first termers at basic training hinges upon the ability of the individual to adjust to his or her environment after leaving the home. Bachelor restrictions are also perceived differently by the teenage airman than by older bachelor airmen who had previously been married.

Another aspect of one's perception of livability is illustrated by the scenario of a satisfied and happy married airman or officer at a sunbelt beach base who receives orders for Minot, ND, about which he knows little other than tall tales (which might be true) and the fact that it is reputed to have low physical livability. In checking with others he is subject either to new rumors, generally negative, or to a "true" report of a colleague who was there briefly on TDY in February and who is more than willing to share his few facts and second hand knowledge.

The next step is to muster positive enthusiasm to break the word to his wife, who is easily influenced by negative rumors in their already shaky marriage and to their teenage daughter, just selected for next year's cheerleader role, and their son who was standby quarterback to a senior last year at high school.

The situation comedy could easily be perceived as a discriminatory tragedy by individual family members and it might encourage some senior people to leave the military service. In truth, feedback studies at Minot show a high rating for social livability despite a physical environment with few amenities.

There is a popular belief that livability is subjective and cannot be quantified. This report holds that is is quantifiable and that military individuals and families already make these judgments. Many use a type of index system before moves, upon arrival, and several years later when readying for departure. For some, it approaches a numerical system (see Airman, March 1977, inside cover story) and it can be used to compare changes in a single base over time or concurrently with a series of bases in a region.

Unfortunately, arbitrary administrative decisions, political factors, adverse community reactions or stressful environmental conditions can seriously lower base livability and start a series of stories moving through the service that may circulate long after the causative factors are gone or resolved.

FAMILY BASE HOUSING LIVABILITY

Early family housing on bases was for senior officers and their families. They were generally large homes and later, smaller units, frequently of brick veneer construction, were added to the small housing cluster. The main expansion took place after WW II, under legislation sponsored by Senators Wherry of Nebraska and Capehart of Indiana. Respective housing clusters constructed on bases were named after the sponsors.

Since 1963, new married housing was authorized by Congress on a project by project basis and these are identified as Appropriated Fund housing. Some were with Title I funds, and since 1973 a number were constructed by local contractors as turnkey housing.

In response to a request by congressional sub-committee the three services undertook an occupant opinion survey in 1972. The Air Force and Navy questionnaire focused on structural details and comfort preferences. The Army questionnaire sought to identify which structural variations should reflect different environments, reasons for living on and off the base, street arrangements, visual

privacy, audio privacy, comparions of military and civilian shopping centers, role of Officer and NCO Clubs, libraries, day care centers and even preferences for diverse house exteriors and roof types.

Greatest user satisfaction was shown with Capehart housing and least was for Wherry housing. The most desired improvements for all housing, ranked in order, were (1) central air conditioning (2) fenced in yards and (3) a 1/2 bath on the first floor of second floor units. Occupants of Wherry housing listed sound-proofing and the need for more interior bulk storage along with air conditioning as their top three priorities.

A 1975 Air Force Housing Study done for AFCEC, Phase II, updated and evaluated family housing to spring 1975. Some of its findings appear in Table V.

T A B L E V

FAMILY BASE HOUSING - SPRING 1975
150,000 UNITS

PERIOD	1950-55	1956-63	1963-1973	1974-1975
NAME	Wherry	Capehart	Appropriated + Title I	Appropriated + Turnkey
TYPES	S.F.Det, Dupl, Apts	S.F.Det, Dupl	S.F.Det, Dupl, row	S.F.Det, Dupl, row
AVERAGE SIZE	900 sq ft	1100 sq ft	1250 sq ft	1370 sq ft
UNIT COST	$10,000	$16,800	$17,500 - $27,500	$30,000
No./%	27,000/ 18%	55,000/37%	60,600/40%	7,400/5%
QUALITY	Marginal	Better design & construction	Acceptable but much variance in projects	Generally marginal but variable

Despite numerous complaints and suggested improvements from questionnaire respondents, over 50% of the airmen surveyed in 1972 and 63% of the officers were satisfied to very satisfied with their base housing. At the other end of the spectrum 27% of the

airmen were dissatisfied to very dissatisfied compared to 21% for the officers. The dissatisfaction levels with Wherry housing ran 12 percentage points higher for both officers and airmen.

In the housing rehabilitation program that started in 1974 the objectives were to upgrade the functional and physical quality of the Wherry units to a level commensurate with the Capehart houses in order to gain increased occupant satisfaction and raise morale. The Building Technology - 1975 study, (Section II-6) felt that the rehabilitation program was heading in the right direction but that it did not show proper coordination and management. It decried the lack of professional design imagination, the lack of design professionals, and the lack of site planning. Though Mr. Meredith at Hq USAF/PRE states, "We must provide Air Force people with the conveniences of life as the civilian sector does for its people,"there is little follow-up and aesthetics and site planning were sacrificed in the rehabilitation effort.

Another independent study of the Air Force Base Housing, completed by the Environmental Planning Institute and submitted on 6 February 1976, states, "the environment provided by the rehabilitated dwelling units rates very poorly in terms of livability" and that it is "not producing a living environment comparable to the private sector." It also attributes part of the shortfall to ignoring the necessary amenities within the site planning. New base family housing tends to be improving despite these rehabilitation problems because of more square footage and design improvements. The assigning of priorities in the remodeling effort is largely based on DoD Instruction Sheet 4165.45, dated 19 January 1972 which set the guidelines to be followed.

There are still numerous problems in Base Housing that affect livability including the following:

- o People in older and smaller housing feel disadvantaged compared to those in newer,larger housing and tend to regard it as a financial inequity
- o People in colder climates indicate that garages are needed rather than carports
- o BCE groups note that many homes with garages tend to use them for storage and then park their cars on the streets
- o People have strong feelings that favor interior pedestrian and bicycle paths connecting homes to schools to avoid using arterials

- o Shortages of interior and exterior storage

- o Too little visual privacy between adjacent or facing homes

- o Too little accoustical privacy

- o A feeling of excessive regimentation

Military families have special incentives for acquisitions. With overseas duty, new cultural exposures and the movement of household goods by the government, storage will always be in short supply.

Extensive literature searches on housing identified a variety of factors from the civilian world but only a few are selected because of their immediate application to base housing.

- o Fences and hedges are perceived as discouraging unwanted guests.

- o Without assigned territory (lot lines or yard allocation for upstairs-downstairs duplexes) there is little sense of privacy and it discourages plan making and execution as well as identity.

- o People tend to become friends with their nearest neighbors, especially if they have young children who play together.

- o Friendships between close neighbors is more important for women than for men especially in one car families or where there is no bus service.

- o The desire to move is based on the present level of satisfaction (push factor) compared to the preceived level of satisfaction to be attained elsewhere (pull factor). The gap between the two represents "stress."

BACHELOR HOUSING

At one time all bachelors lived in barracks or in the BOQ. Now, officers with the rank of captain or above may draw allowances and live off base. Single enlisted and junior officers are permitted to live off base without allowances. Some choose this for greater freedom and privacy. Higher grade officers stay at the BOQ temporarily or because they dislike commuting and prefer to be closer to their work and certain base facilities. Enlisted quarters are dominated by first or second term airmen and women, usually in different buildings or in separate wings of the same structure. The tendency to segregate men and women is a relict behavior pattern.

To grasp livability concerns in bachelor quarters visits to individual rooms and day rooms were made in the company of billeting officers, first sergeants or their representatives to observe bedrooms, day rooms, layouts, acquisitions, and storage facilities and to interview individuals and supervisors. A second visit, with a base photographer,resulted in a narrative and slide materials which were turned over to the head of the Environmental Planning Directorate.

The slides were taken to show problems of interior and exterior storage, the high entry privacy but low window privacy, attempts to achieve individualization and privacy in rooms by dividers, scheduling, and the use of headsets for accoustical privacy.

Other elements noted were storage or security problems with hibachis, bicycles and motorcycles, the growing use of plants, the popularity of stereo sets with up to eight speakers in a two person room, tapedecks, record players and personal television sets instead of just one in the day room.

The minimum square footage per person in bachelor airman quarters is spelled out in several DoD and AF regulations. The Air Force reluctantly maintains a maximum of two persons per room because of pressures from Congress and other services that require three or four persons per room. In a few instances basic trainees may still live in open bay barracks - a situation that fosters a sense of community, necessary for new recruits despite the lack of privacy.

The,older one story dormitories have more square footage per individual than the newer two and three story buildings. This occurs at a time when airmen, reflecting the rest of society, acquire many new leisure time oriented items such as stereo sets, tape decks, bicycles or motorcycles, TV sets, coffee makers, books, plants, etc.

All officer's quarters seen and almost every airman's room visited had a refrigerator. Many contained beverages, crackers or other sweets that would attract roaches yet toasters were prohibited on the grounds that it constituted food preparation and might attract roaches. Another objection was raised on the grounds of energy consumption but this question was not raised concerning the other electrical appliances.

Several questions raised as a result of this field excursion were:

- Should bachelor housing move to minimum square footages when life styles involve numerous leisure time acquisitions?

o Why enlarge family units while retaining minimum size bachelor quarters?

o Should external rental storage facilities be built on bases for bicycles, motorcycles or smaller items?

o Should special provisions be designed to accommodate bicycles and motorcycles?

o Should electric meters be installed and charges made to encourage energy conservation?

o Should more area be provided for second termers assuming more time to acquire possessions?

o Should rooms retain double occupancy for first termers that convert to single rooms for second and third termers on a space available basis?

Other questions that produce or will produce dormitory problems are the seeming parental concern for morality by inspections, hours for women, prohibiting overnight guests of the opposite sex, and the introduced inequity from more liberal square footage allowances, shower doors and kitchen privileges for women. Many of these issues were resolved since 1972 in college dormitories but they remain irritants in the military.

Using a framework of comfort, sense of community, convenience, privacy and opportunities for individualization, a livability index grading system was prepared to be applied to bachelor dormitories (Table VI) based on a maximum 5-point system but evaluated by a joint team representing the Base Civil Engineering staff, the billeting office and occupants. Individual variations can be made to increase relevance to a specific type of building on a particular base and then evaluated against readiness, livability and productivity.

T A B L E VI

LIVABILITY INDEX - BACHELOR HOUSING

Full compliance - 5 points: fewer points with partial compliance.

- o Design appropriate to region and site
- o Senior officer or Dormitory Chief maintains standard of cleanliness, noise, personal conduct that respects rights and privacy of all occupants
- o Rooms and halls - carpeted,soundproofed
- o Areas provided for guests and group interaction
- o Provisions for privacy and personal space
- o Billeting office and users encouraged to provide feedback on design, maintenance, administrative practices

- o Occupants encouraged to improve room decor
- o Contains some moveable furniture
- o Variety and harmony in colors, drapes, carpets
- o Internal and external storage
- o Nearby picnic and hibachi areas
- o Phones on each floor
- o Windows for outside contact
- o Occupants perceive low density
- o Sense of community encourages

In a 1969 study, white college administrators were asked what they thought the goal and purpose of a dormitory should be. They listed togetherness, intimacy, and informal life style. In response to the same question, students indicated that they wanted privacy. There is value in pursuing this question within the military to match perceptions between users, administrators and the designers.

WORK LIVABILITY

For some the objective is less work, shorter hours, more people to share the work, ever increasing benefits, more disposable income, faster advancement and more discretionary leisure time. The ultimate for them is no work. However, most people in America are still committed to a personal work ethic. They want something worthwhile to do, clear work goals, resource choices, feedback, an opportunity to communicate with the rest of the system, more training and responsibility, visibility, recognition and the opportunity to build a reputation.[6]

The prescription to work remains high. Americans want increasing pleasures and satisfactions and a sense of worth from work, matching the evolution of rising expectations in housing and the opportunity to acquire more security, goods, services and experiences.

The work design for a military organization is for total effectiveness during wartime operations. Peacetime creates problems in keeping up work morale and in producing a positive feeling of capabilities, accomplishament and public service. At Air Force bases, there seems to be a difference between the top priority flight line operations and lower priority support functions.

In the civilian industrial section, some work livability indexes have been based on (1) productivity (2) worker turnover (3) absenteeism (rates, and patterns by subunits) (4) quality of the workmanship and (5) job satisfaction.

The principles behind these factors also apply to military installations but work indexing an entire base would not be meaningful.

Indexing is complicated by the civilian-military mix, the airman, NCO, junior and senior officer mix in a rank and job rated environment, seemingly excessive military turnover for training and reassignment, and changing interpersonal relations that never seem to stablize. There is concern about getting clear goals and subgoals, fragmentation of work assignments and responsibilities.

If questionnaires and personal observations at four air bases are valid, it appears that many feel that only limited job satisfaction, that productivity varies greatly between buildings, offices and functions, that significant numbers of married enlisted personnel who live off base moonlight at second civilian jobs, that many regard it to be more important to arrive for work on time and to be present the entire day than to point to significant daily work accomplishments, that the main work objectives are high personal ratings and to be ready in case of an emergency.

WORK AREAS AND CROWDING

A common work complaint in support areas is that of crowding, expecially if other individuals or units seem to operate with a larger square footage. Probably more disturbing than the sense of shrinking space, is the possible related self image of shrinking importance and the new interpersonal relationships that have to be established and resolved by open work areas, partitions, audio and visual privacy mechanisms,and finally, satisfactory clarification of the status heirarchy role with the new intruder(s). It is not unusual for operations, at least on a cyclical pattern, to be in a state of turmoil because of realignments, mission changes, endless reorganizations, space relocations, new temporary people, in or away on short stays, leaves, etc. Several studies indicate that people working in more crowded space develop adaptive behaviors and that the crowding merely intensifies normal reactions. Carpeting, light, and other atmosphere controls can help offset crowding problems. Creative and high concentration activities require more privacy and quieter environments.

Once physical comforts are provided and job roles clearly established, the major elements that lead to alienation, boredom, depression, loss of interest and dissatisfaction are keyed to monotony, confused personal goals, and interpersonal relations on the job. In this environment, minor work-related irritations grow with time and the work arenas are used for venting hostilities generated in non work environments - such as domestic, social, financial, alcohol, commuting, children, cars and parking.

Conversely work problems maybe vented at home or recreation type pursuits. Supervisors are expected to be sensitive to detect problems, to listen or refer personnel to helping units. However, supervisors may not recognize when they, themselves, have the problems.

Based on numerous polls, workers want "meaningful work" ahead of higher pay, more benefits, new facilities, shorter hours, promotion and job security. This finding is in direct conflict with conventional wisdom which has long held that increased pay and benefits are uppermost in the minds of most job holders. The key to a better workplace is in the job itself. Despite more benefits and money, a "dumb" job is still a "dumb" job.

Goals and guidelines should aim to:

1. create opportunities for people to use their minds and training on the job.

2. replace heavy handed bossing with greater individual responsibility for getting the job done.

3. invite employees at all levels to provide feedback to the decision making process.

4. convince employees that you value their good recommendations by fighting for them at the next higher eschelon of authority.

5. allow individuals greater control over their time as long as it insures greater or equal productivity.

6. redesign, eliminate or consolidate jobs that have become monotonous and meaningless.

7. insure reasonable treatment that reflects human respect, recognizes individual needs, aspirations and capabilities

8. advocate the support services for civilians as well as military that reduces stresses and increases job efficiency.

The paternalistic attitude evident toward bachelors in housing has only minor residual carryovers to the work environment. Excuses from work to transport immediate family members for medical appointments or to airports are generally approved pro forma but only with reservation were it to apply to boy friend or girl friend situations.

There is an interface between work, recreation and physical fitness. Military individuals in sedentary work assignments are also expected to meet selected minimal service wide physical fitness standards as a basis for retention in the service.

Obesity is a major cause for failures, yet base after base spawns hundreds of vending machines in work areas specializing in high calorie, low nutritional items dominated by the ubiquitous pop machine. Snack bars favor food high in chloresterol - bacon, sausage, eggs, and even whole milk are standard fare. No serious effort seems evident to support a health and nutritional concern that matches that of physical fitness. Therefore a major objective in recreation programs and facilities is to recreate the physical fitness that other actions tend to undo. Isn't nutritional training as important as human relations?

There is no single approach in the way individuals view their specific work assignment. The pilots convey a feeling of enjoyment and satisfaction though not always with specific flight destinations, early morning training flights nor the extended departures from their families. The further one departs from the flight line, the more the work seems to be regarded of lesser importance, like just another government agency job and less like a military career or calling.

SUMMARY AND CONCLUSIONS

Livability is still popularly perceived to be either (a) a subjective highly personalized value system, (b) more justifiable benefits for "us" but not for others, and (c) a "buzz" word to be used for current or future gains tied to more leisure time options, better housing, less work and fewer responsibilities.

This study has shown a number of common elements in base livability especially in housing and work areas despite pronounced differences between treatment of bachelors and married personnel, on and off base restrictions, flight line and support groups, civilian and military personnel.

Individuals and groups are always making comparative mental livability indexes. More objective means are possible and need development.

The Air Force provides important livability benefits that help attract men and women to service careers. Second generation Air Force families provide positive hard test measures of perceived livability and desirability.

Livability is an important people sensitive consideration in all operational and planning processes and it involves many groups on a base. It should be incorporated in comprehensive planning but be carefully weighed against other Air Force objectives including readiness, productivity, a sense of community and mission effectiveness.

In hindsight the study recognized that livability has several measurable components and that leisure time activity options, the physical environment, housing and work areas are key ingredients. Though individual livability is comparative to prior personal experiences, collective group perceptions of base livability can be measured indirectly by reenlistment rates, complaints, requests for transfers to or from specific bases, and several key attitudinal elements. Livability varies between bases, domestic and foreign, probably between commands and it may vary with the size and distance of nearby cities and the nature and variety of physical environmental options. These last named items are recommended for further study.

FOOTNOTES

1. The Quality of Life Concept: A Potential New Tool for Decision-Makers, E.P.A. Office of Research and Monitoring Environmental Studies Division, Washington, D.C.,Chapter 5, pp. 1-80, 1972 (?)

2. Hornback, et al, Studies in Environment, Vol. II, Quality of Life, Washington Environmental Research Center, EPA Grant No. 801473, November 1973,U.S. Superintendent of Documents, Washington, D.C., pp. 36-38.

3. Hutcheson, R. G., Jr., "The Churches and the Chaplaincy," John Knox Press, Atlanta, 1975, p. 93.

4. Henriksen, Lester H., and Gary D. Vest,"Mission Effectiveness and Base Livability," Air Force Engineering and Services Quarterly, November 1976, pp. 2-6.

5. Altman, I., "Privacy: A Conceptual Analysis," Environment and Behavior, Vol. 8, No. 1, March 1976, pp. 7-30.

6. Johnson, J. Richard, "Job Livability," Air Force Engineering and Services Quarterly, November 1976, p. 5.

OTHER REFERENCES

o *Tri-Service 1972 Housing Occupant Surveys*

o Executive Summary, Air Force Housing and Materials Analysis, Environmental Planning Institute, Sausalito, CA 1976

o Phase II Air Force Housing Study,for AFCEC, by Building Technology, Inc., Silver Springs, MD 1975

o Environment and Behavior - numerous articles in this journal

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

NONDESTRUCTIVE PAVEMENT EVALUATION

Prepared by:	Braja M. Das
Academic Rank:	Associate Professor
Department and University:	Department of Civil Engineering, South Dakota State University
Assignment:	Tyndall AFB. Air Force Civil Engineering Center, Directorate of Engineering Materials, Pavement Evaluation Division
USAF Research Colleague:	Major James I. Clark
Date:	August 5, 1977
Contract No:	F44620-75-C-0031

ABSTRACT

Nondestructive Pavement Evaluation

by

B. M. Das

Research works have been in progress for about ten years to develop a compatible pavement evaluation procedure for airfields based on nondestructive tests. A successful nondestructive pavement evaluation technique will reduce the time of closure of various airfield facilities which is needed to conduct destructive test required for conventional pavement evaluation.

This study provides a comparison of projected pavement lives of several airfield features estimated by nondestructive and destructive pavement evaluation procedures. For aircraft and gross load on similar pavement section, the nondestructive evaluation procedure yields higher number of allowable operations as compared to that obtained by the destructive test evaluation technique. The wide variations may be due to inaccurate estimation of the elastic modulus of subgrade.

ACKNOWLEDGMENT

The author is grateful to the Air Force Systems Command, Office of Scientific Research; the Engineering Materials Directorate of the Air Force Civil Engineering Center and the American Society for Engineering Education for the technical support of this research effort. Administrative support was provided by Auburn University, Alabama.

The author also wishes to acknowledge the support of his USAF Research Colleague, Major J. I. Clark. It has been a pleasure to work with Capt D. Burk, Capt T. Black and Capt D. Artman. In particular, thanks are extended to Lt Col George D. Ballentine for his assistance during the period of this study.

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

AGL	Allowable gross load
a	The ratio of Rayleigh wave velocity to shear wave velocity
d	Distance between compared accelerometers
E	Young's modulus
E_{cor}	Corrected Young's modulus
f	Flexural strength of concrete
G	Shear modulus
g	Acceleration due to gravity
GI	Aircraft group index
PI	Plasticity index
P/C	Pass-per-coverage ratio
S_r	Degree of saturation
t	Tensile strength of concrete
T	Thickness of a pavement layer
v	Phase velocity
V_R	Rayleigh wave velocity
V_S	Shear wave velocity
γ	Unit weight of material
ϕ	Phase angle
λ	Wave length
ν	Poisson's ratio

INTRODUCTION

One of the important responsibilities of the Air Force Civil Engineering Center is to evaluate the load carrying capacities and the remaining service lives of airfield pavements. The pavement evaluation procedure presently adopted requires destructive testing. This requires the closure of runways and taxiways for extended periods of time for completion of the in-situ tests and collection of samples for laboratory testing.

In order to improve the airfield pavement evaluation technique, the Air Force has been involved in developing a compatible nondestructive test procedure since 1967. The essence of this program is based on determination of the elastic properties i.e. shear modulus and Young's modulus of each of the layers which constitute a given pavement section by means of vibratory testing. Subsequently, by using these results a computerized structural analysis is made to determine the allowable number of prefailure operations for a given aircraft in that section of the airfield pavement. A historical review of the nondestructive pavement evaluation procedure has been presented by Nielsen and Baird (Ref 1). The present state-of-the-art for this technique has reached a considerable degree of sophistication.

The purpose of the present study is to compare the projected pavement life of various features of some airfields obtained from the analysis of destructive tests with that obtained from the analysis of nondestructive tests.

PAVEMENT EVALUATION PROCEDURE BASED ON DESTRUCTIVE TESTS

This is the present technique for evaluation of the allowable gross loadings for airfield pavement features. The procedure for evaluation of flexible and rigid pavements are given in Refs. 2 and 3 respectively. For evaluation purposes, various aircrafts are placed under a number of group indices (GI) as shown in table 1. This classification is based on the number of wheels, wheel configurations and tire contact areas of the aircraft. For any given airfield feature, the pavement evaluation study yields an estimate of the allowable gross loads (AGL) for each of the group indices under operational categories such as capacity, full, minimum and emergency. The operation categories mentioned above are based on the coverage levels given in table 2. By definition, a coverage is said to occur when all points of the pavement surface within the traffic lane have been subjected to one application of maximum stress by the design aircraft.

The important field results needed for determination of AGLs for a given section of airfield are given below.

(a) Flexible pavement:

(i) Thickness of each layer constituting the pavement section and

(ii) CBR values of all soil layers

(b) Rigid Pavement:

(i) Thickness of the layers of the pavement section

(ii) Flexural strength of Portland cement concrete surface layer and

(iii) The modulus of subgrade reaction of underlying layers.

AIRCRAFT GROUP INDICES

SINGLE WHEEL			MULTIPLE WHEEL								
TRICYCLE			TRICYCLE						BICYCLE	TRICYCLE	DOUB TRICY
1	2	3	4	5	6	7	8	9	10	11	12
C-123	A-7	F-111	C-7	C-130	EC-121		C-135	C-5	B-52	B-1	E-4
	A-10	FB-111	C-9		KC-97		C-141				
	EB-57		C-54				E-3A				
	F-4		C-131				KC-135				
	F-5		C-140				VC-137				
	F-15		T-29								
	F-16		T-43								
	F-100										
	F-101										
	F-102										
	F-105										
	F-106										
	T-33										
	T-38										
	T-39										

TABLE 2

COVERAGES FOR VARIOUS AIRCRAFT OPERATION CATEGORIES

Operation Category	Number of Coverages: Channelized Section	Number of Coverages: Nonchannelized Section
Capacity	25,000 for all aircraft except B-52; 10,000 for B-52	5,000
Full	5000	1,000
Minimum	1000	200
Emergency	200	40

PAVEMENT EVALUATION PROCEDURE BASED ON NONDESTRUCTIVE TESTS

The nondestructive pavement tests used for evaluation purposes in this study were conducted with the type of equipment developed at Civil Engineering Research Facility, University of New Mexico. All the components of the test equipment are placed in a van 8 ft wide and 35 ft long. For the nondestructive tests, a dynamic load of 1000 lb. which can be varied sinusoidally is applied by a vibrator on the surface of the pavement and frequency sweep between 10Hz to 3500 Hz is conducted. The vertical acceleration of the pavement at selected distances from the applied load is measured by accelerometers which are epoxied to the surface of the pavement. A phase computer is used to determine the phase angle from the signals recorded between any two accelerometers.

The phase angle/frequency plots obtained from field tests are then reduced to obtain plottings of wave length against phase velocity which are referred to as the dispersion curves. The relations for the wavelength and phase velocity are as follows:

$$\lambda = \frac{360d}{\phi} \quad (1)$$

$$v = f\lambda \quad (2)$$

where, d = distance between the compared accelerometers in/ft.

ϕ = phase angle in dlg.

f = frequency in Hz

λ = wavelength in ft.

v = phase velocity in ft/sec.

The dispersion curve mentioned above can be obtained by the use of a computer program, NDTPLOT, which does the calculation of the values of λ and the corresponding v and completes plotting of the dispersion curve.

The peak wave velocity obtained from the dispersion curve gives the Rayleigh wave velocity, V_R, for the surface layer. The shear wave velocities in the underlying layers (i.e. base/subbase, subgrade) are obtained from the ordinates corresponding to the break points of the curve. The shear wave velocity, V_S, of the surface layer can be determined from the relation,

$$V_S = \frac{V_R}{a} \quad (3)$$

where 'a' is a function of Poission's ratio, ν, of the material and it's theoretical values are:

$a = 0.875$ for $\nu = 0$ and

$a = 0.955$ for $\nu = 0.5$.

Making linear interpolation, the following values of 'a' have been determined for calculations in this study:

a = 0.899 for pavements with concrete surface layer and

a = 0.911 for pavements with asphaltic concrete surface layer.

With the shear wave velocities known, the Young's modulus for each layer of a given pavement section can be determined from the following theoretical relations:

$$G = \frac{V_s^2}{144} \frac{\gamma}{g} \quad (4)$$

$$E = 2\,(1 + \nu)\,G \quad (5)$$

where, G = shear modulus of the layer in lb/in^2

E = Young's modulus of the layer in lb/in^2

g = acceleration due to gravity = 32.2 ft/sec^2

V_s = shear wave velocity in ft/sec, and

γ = unit weight of the material in the layer in lb/ft^3

The average approximate values of γ and ν for different materials encountered in an airfield pavement section are given in table 3.

TABLE 3

REPRESENTATIVE VALUES OF UNIT WEIGHT AND POISSON'S RATIO

Material	γ (lb/ft^3)	ν
Concrete	145	0.15
Asphalt	145	0.43
Base Course	120	0.25
Subgrade	110	0.43

Comparison of field tests have shown that the Young's moduli obtained from the above calculations yield somewhat higher values for aphaltic concrete surface layers, base/subbase layers and subgrades. Hence for actual pavement performance evaluation, corrections are made to obtain representatives values of E. The present recommended correction procedure is given in table 4.

TABLE 4

CORRECTED YOUNG'S MODULUS, *

Layer	E_{cor}
Surface	
Concrete	$E_{(concrete)}$
Asphaltic concrete	$\frac{E_{(asphalt)}}{2}$
Base/Subbase with	
Concrete Surface	$\frac{E^2_{(base/subbase)}}{E_{(concrete)}}$
AC Surface	$\frac{E_{(Base/subbase)}}{2}$
AC/concrete surface	$\frac{E_{(Base/subbase)}}{2}$
Subgrade with	
Concrete Surface	$\frac{E_{(subgrade)}}{2}$
AC Surface	$\frac{E(subgrade)}{2}$
AC/PCC Surface	$\frac{E_{(subgrade)}}{2}$

*Ref 1.

Note: $E_{(concrete)}$, $E_{(asphalt)}$, $E_{(base)}$, $E_{(subgrade)}$, correspond to Young's moduli of the materials represented by its subscripts obtained by using eq. 5.

Ultimately, with the above calculations completed, the pavement evaluation is made by a computer program known as PREDICT. It is a nonlinear finite element program which performs the structural analysis of the pavements in which the aircraft wheel loads are represented by Fourier Series and is capable of treating single and multiwheel landing gears. In order to obtain the number of passes that an aircraft can operate on a given pavement section before fatigue failure occurs. The following are used as input to the PREDICT code:

(a) Name of the aircraft (At the present time, the aircraft in PREDICT code are B-1, B-52, B-57, Boeing 747, C-5, C-9A, C-130, C-141, F-15, F-16, F-105, F-111, FB-111A, KC-97, KC-135 and T-39).

(b) Number of layers in the pavement section,

(c) Thickness, T, of Surface layer, base course, subbase (if present) and subgrade. The thickness of subgrade is given by:

$$T_{\text{Subgrade}}\ (\text{in}) = 144 \text{ in.} - T_{(\text{surface + base + subbase})}\ (\text{in}) \qquad (6)$$

(d) E_{cor} for all the layers.

(e) Tensile strength of concrete layer, t (in psi), if present.

(f) Poisson's ratio for each layer.

(g) Degree of saturation, S_r, for the soil layers

(h) Void ratio, e, for the soil layers.

(i) Plasticity index, PI, for the soil layers.

(j) Nature of traffic on the pavement i.e. channelized or nonchannelized.

The failure criteria used for the destructive test evaluation theory are not the same as those used for the nondestructive test pavement evaluation procedure. The criteria used for destructive tests are based on the development of the first crack in the pavement and derived from large scale field tests conducted over the years. The basic assumptions for failure used for nondestructive pavement analysis are given in reference 1. There is a need for a comprehensive study to see

if the assumption of "failure" used for nondestructive pavement evaluation theory is acceptable.

PROCEDURE FOR COMPARISON OF PROJECTED PAVEMENT LIFE BASED ON DESTRUCTIVE TESTS TO THAT BASED ON NONDESTRUCTIVE TESTS.

In order to make a reasonable comparsion to see how well the projected pavement life based on destructive tests compares with that based on nondestructive tests, the following procedure has been adopted in this study. It is presented in a step by step manner for easier understanding. All comparisons presented here are made for capacity category operation of aircraft.

(a) Review the AGLs for a given feature of an airfield for capacity category operation obtained in the pavement evaluation report based on destructive tests.

(b) Determine the group indexes of the aircraft in the PREDICT code.

(c) Compare the maximum gross load (Table 5, column 2) of an aircraft considered in step-b with the AGL obtained for its corresponding group index in step-a. If the maximum gross load of an aircraft is less than or equal to the AGL for the corresponding GI, it may be selected as a case for the comparison study.

(d) For an aircraft selected in step-c, determine the number of permissible operations as predicted by the destructive evaluation technique. Or,

$$\text{No. of operation} = \underset{\substack{\uparrow \\ \text{Table 2}}}{\text{No. of coverage for Capacity Category}} \times \underset{\substack{\uparrow \\ \text{Table 5, Col. 4 \& 5}}}{\frac{P}{C}} \qquad (7)$$

P/C = pass-per-coverage ratio.

(e) Determine the values of E for various layers of the pavement (for the feature under consideration) from nondestructive test(s). If more than one test has been conducted in that feature of the airfield, the moduli can be calculated from respresentative values of v_s for the layers.

(f) Correct the values of E obtained in step-e by using table 4.

(g) Determine the tensile strength, t, of the concrete layer (if present) from tensile splitting tests of four and

six inch diameter concrete cores.

TABLE 5

Gross Weight, Main Gear Wheel Load
and Pass-Per-Coverage Ratio for Some Aircraft

Aircraft	Maximum Takeoff Gross Weight* (KIPS)	Maximum Main Gear Wheel Load* (KIPS)	P/C ** Channelized	P/C ** Non-Channelized
(1)	(2)	(3)	(4)	(5)
B-52H	488.0	67.1	1.63	2.0
B-57B	58.8	27.7	6.47	12.83
C-5A	769.0	30.2	0.81***	1.10***
C-130E	175.0	41.9	2.09***	4.05***
C-141A	316.6	37.4	1.72***	3.17***
F-105F	54.6	23.4	10.9	21.9
F-111A	98.6	47.0	4.92	9.8
KC-97G	187.0	44.5	3.41	6.11
KC-135A	300.8	35.5	1.68***	3.03***

* Reference 4.

** Reference 5.

*** Pass-per-coverage ratio for rigid pavement is equal to twice the value shown.

(h) Determine the main gear wheel load, WLOAD, for the aircraft by using the relation,

$$WLOAD = \frac{AGL}{\text{(Max gross load of aircraft)}} \times \text{[Max main gear wheel load]}$$

(Max gross load of aircraft): Table 5, col. 2

[Max main gear wheel load]: Table 5, col. 3

(i) Input the following to the program, PREDICT.

Name of aircraft

Nature of traffic (channelized or nonchannelized)

WLOAD (Step-h)

t (Step-g)

Number of Pavement layers

Poisson's ratio for each layer (table 3)
Plasticity index (PI) for each soil layer.

Void ratio, e, for each soil layer (actual value if known or about 0.22 for base course and 0.45 for subgrade).

Degree of saturation, S_r, for each soil layer (actual value if known or about 80%)

(j) PREDICT gives the number of allowable passes for the aircraft on the pavement section.

The number of passes obtained in step-j by the non-destructive evaluation technique can now be compared with the number of passes obtained in step-d by the destructive evaluation technique. It may be noted that the number of passes obtained by the above two methods are for the same aircraft and gross load.

A flow chart for the comparison procedure (step-d through j) is given in figure 1.

RESULTS OF COMPARSION AND DISCUSSION:

The results of comparison of pavement life for several features of four airfields are given in table 6. It is observed that in all cases, the number of permissible operations calculated on the basis of nondestructive test evaluation procedure is appreciably higher than that predicted on the basis of the conventional destructive test evaluation procedure. Based on these limited results, the correlation coefficient was found to be 0.29.

The higher number of operations predicted by nondestructive evaluation procedure can possibly be due to the unusually high value of V_s and thus E_{cor} for subgrade obtained from the analysis of nondestructive test results. For example, the values of E_{cor} for subgrades (unified soil classification - CL) obtained from the nondestructive tests of Carswell AFB were in the range of 20×10^3 to 120×10^3 psi. Similarly in the case of Dyess AFB, the E_{cor} for CL type subgrades ranged from 55×10^3 to 135×10^3 psi. The thicknesses of the concrete layers in the pavement for the above cases varied between 15 to 26 inches. However, the usual range of Corrected Young's modulus should be about 3×10^3 to 10×10^3 psi. This discrepancy can possibly be attributed to the dominance of the stress wave velocities in thick concrete layers. This makes the validity of empirical modulus correction procedures of subgrades given in table 4 somewhat questionable. It may be pointed out the above empirical correction procedure has been developed from the limited results of experiments conducted at Cannon AFB, New Mexico. At Cannon AFB, the thickness of concrete surface layer of pavement varies from about 8 in to about 14 in with an average thickness of about 11 in. Thus, it may be speculated that the empirical procedure of 50% reduction of calculated Young's modulus to arrive at E_{cor} may not hold good in all cases.

CONCLUSIONS AND RECOMMENDATION FOR FUTURE RESEARCH:

To meet the primary objective of this research effort the following recommendations are made:

(1) The nondestructive pavement evaluation procedure is theoretically sound. However, in its present form, it appears to give overly optimistic results and does not seem to be ready for general use. The main problem lie in the inability of determining the Young's moduli of subgrades where surface layers are concrete with thickness of about 10 in. or more and the variation between the failure criteria used in the destructive versus the nondestructive methodologies.

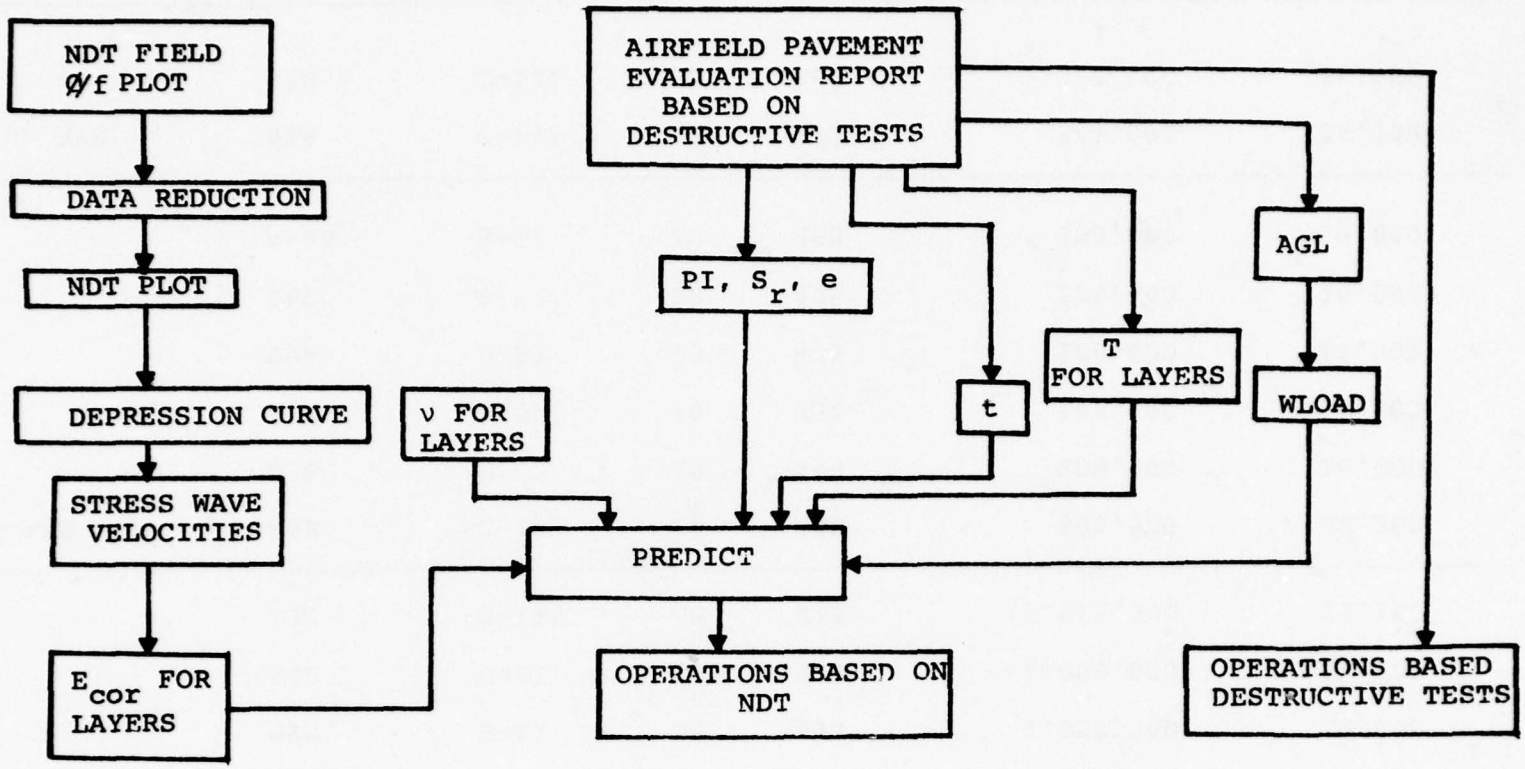


Figure 1. Flow chart for comparsion of pavement life based on destructive tests to that based on nondestructive tests.

TABLE 6. RESULTS OF COMPARISON STUDY

Air Force Base (1)	Feature (2)	Aircraft (3)	GI (4)	AGL (KIPS) (5)	Operations Calculated from: Nondestructive test results (6)	Operations Calculated from: Destructive test results (7)
Carswell AFB	R2B	B-52	10	330	815,000	16,300
	T9C	B-52	10	295	1,000,000	10,000
	R4C	B-52	10	420	1,000,000	10,000
	A3C	KC-135	8	215	1,515,000	15,150
Dyess AFB	A4B	KC-135	8	240	606,000	30,300
	A5B	B-52	10	290	200,000	10,000
	R1A	B-52	10	350	163,000	16,300
	T9A	B-52	10	475	163,000	16,300
	R6C	B-52	10	405	200,000	10,000
	A-6B	B-52	10	350	200,000	10,000
Keesler AFB	R1A	C-141	8	225	344,000	86,000
	R4B	C-141	8	265	344,000	86,000

TABLE 6. (CONTINUED)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Davis Monthan AFB	R1A	KC-97	6	180	341,000	85,250
	R1A	KC-135	8	280	336,000	84,000
	R1A	B-52	10	230	163,000	16,300
	R2A	F-105	2	40	1,296,427	272,500
	R2A	F-111	3	80	520,857	123,000
	R2A	B-52	10	255	35,980	16,300
	R3A	F-111	3	85	670,125	123,000

Further investigation of Young's Moduli can be accomplished by conducting nondestructive tests on pavements having similar subgrades but varying thickness of concrete surface layers. This should provide information regarding the dependency of the empirical modulus correction factor on the thickness of the concrete layer.

(2) Nondestructive tests have been performed in several other Air Force Bases such as Shaw AFB, Blytheville, AFB, Holloman AFB and Mather AFB. Using these results and the information available from destructive test evaluations, comparison of pavement lives such as those presented in this report should be made.

(3) More research needs to be done to develop acceptable failure criteria for use in the nondestructive pavement evaluation.

(4) It may be desirable to increase the magnitude of dynamic load applied to the pavement surface during the nondestructive test. This could give better results.

REFERENCES

1. Nielsen, J. P. and Baird, G.T., "Pavement Evaluation System," AFCEC-TR-76-28, Oct 1976, Air Force Civil Engineering Center, Tyndall Air Force Base, Florida.

2. "Flexible Airfield Pavement Evaluation," AFM 84-24, Chap 2, Oct 68, Department of Army and Air Force.

3. "Rigid Airfield Pavement Evaluation," AFM 88-24, Chap 3, Feb 1959, Department of the Air Force.

4. Hay, D. R., "Aircraft characteristics for Airfield Pavement Design and Evaluation," Technical Report No. AFWL-TR-69-54, Oct 1969, Kirtland Air Force Base NM.

5. Brown, D. N. and Thompson, O. O., "Lateral Distribution of Aircraft Traffic," Miscellaneous Paper S-73-56, July 1973, US Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

EFFECT OF COOLING WATER SPRAY ON TURBINE ENGINE TEST CELL EMISSIONS

Prepared by:	Charles Springer, Ph D
Academic Rank:	Professor
Department and University:	Department of Chemical Engineering University of Arkansas
Assignment:	Tyndall AFB Det 1, ADTC (Civil and Environmental Engineering Development Office) Environics Directorate Air Quality Research Div
USAF Research Colleague:	Maj Peter S. Daley
Date:	August 5, 1977
Contract No.:	F44620-75-C-0031

EFFECT OF COOLING WATER SPRAY ON TURBINE ENGINE TEST CELL EMISSIONS

by

Charles Springer

ABSTRACT

The processes occuring in turbine engine test cells were analyzed to determine the effects of cooling water spray on the emissions. A predictive model was developed to compute the amount of gaseous and particulate emissions as well as plume visibility as functions of engine and cell operating parameters.

It was found that very little benefit may be derived from the use of a quantity of water in excess of that required to cool the exhaust to a temperature which is safe for the equipment.

ACKNOWLEDGEMENT

The author is indebted to the personnel of the Air Quality Research Division of the Civil and Environmental Engineering Development Office for their friendly and helpful attitude which provided a cordial and relaxed working environment. Special thanks for assistance and guidance are due to Maj Peter Daley and to Maj Bradford Grems. It has been a pleasurable and profitable experience working with the very capable personnel of the Air Quality Research Division.

The author is grateful to the Air Force Office of Scientific Research for support of this project, and to Mr Fred O'Brien for his capable administration of the undertaking.

NOMENCLATURE

A	Area
A_x	Cross sectional area
B_o	Background light (Brightness)
B_s	Scattered light (Brightness)
C_p	Heat capacity (Energy/mass Δtemp)
D	Diameter
D_v	Diffusivity (Area/time)
E	Emission factor
F	Fuel flow rate (mass)
F/A	Fuel to air ratio (mass)
H	Specific enthalpy (energy/mass)
I	Impacting flux (particles/area - time)
L	Length (of equipment)
M	Molecular wt
N	Particle population density (particles/volume)
N_a	Mass transfer flux (moles/area - time)
P	Pressure
P^*	Vapor pressure
Q_p	Heating value of fuel (energy/mass)
R	Universal gas constant
R_d	Dilution ratio - dilution air to exhaust, mass ratio
Re	Reynolds No
Rn	Ringelmann No

SN	Smoke No
S_p	Specific particle extinction (area/mass)
T	Temperature
T_o	Ambient temperature
T_s	Saturation temperature
T_r	Plume transmittance (Percent)
U	Velocity
U^*	Eddy velocity
V	Volume or volume/time
W	Water vapor quantity (mole or mole/time), mass of sample for smoke no.
We	Weber No
c	Concentration (mole/volume)
g_c	Newton's Law conversion factor
h	Humidity of ambient air (mole vapor/mole dry air)
m	Mass or mass/time
n	No of particles or no of particles/time
r	Radius (particle or droplet)
t	Time
x	Distance or length (path length)
x_i	Stopping distance
y	Mole fraction
Y	Mole ratio

α	H/C (atomic) ratio for fuel
δ	Surface tension
ρ	Density
μ	Absolute viscosity
ν	Kinematic viscosity
ω	Particulate (or droplet) loading (mass/volume)

Subscripts

a	air
d	diffuser, deposited
e	exhaust
g	gas
I	initial
j	component index
o	emitted or exhausted to the atmosphere; also ambient
p	particle (or droplet)
s	stack
t	throat
v	vapor (water)
w	liquid (water)

EFFECTS OF COOLING WATER SPRAY ON ENGINE TEST CELL EMISSIONS

1. INTRODUCTION

Statement of the Problem

The Department of Defense agencies have demonstrated leadership in establishing and maintaining procedures to minimize adverse environmental consequences of various military activities consistent with maintaining an adequate national defense capability and reasonable cost effectiveness. In keeping with this posture, the Air Force is seeking to find operational procedures for turbine engine test cells which will minimize the environmental effects of engine testing.

Air Force and Navy engine test facilities have occasionally been cited by local civil authorities for violation of visible emission standards for stationary sources. The present effort was undertaken to determine what effect the practice of injecting water to cool the exhaust has on the visible as well as gaseous emissions, and to determine if these emissions can be minimized by careful control of the water injection rate.

The Nature of Turbine Engine Emissions

The air polluting emissions from turbine engines consist of carbon monoxide (CO), unburned hydrocarbons (HC), oxides of nitrogen (NO_x), particulate matter and oxides of sulfur (SO_x). The relative emission rates of these pollutants varies considerably depending not only upon the engine model under test, but the operational mode as well.

CO emissions are the highest in the idle mode, and decrease sharply, generally by at least a factor of ten at the higher power levels. CO emission measurements as reported in Ref 1 show variations from almost 200 lb/hr to less than two percent of that value. In general, test cell emissions of CO are not considered to be significant, (Ref 1, 2) and ground level CO concentrations near test cell operations are ordinarily undetectable.

Unburned hydrocarbons are also more prevalent at lower power modes, and show variations similar to CO emissions. In the usual case, HC emissions are concentrated in the lower parts of this range and are less than 20 lb/hr. The HC emissions may not all arise from the engine itself, but can also result from evaporation of fuel which may have been spilled in the test cell. Some actual measurements have shown the cell emissions of HC to be from two to four times as much as the engine emissions, presumably because of the spilled fuel

(Ref 3). Random variables of this sort make generalizations difficult. Naugle, et. al. (Ref 4) show that hydrocarbons are the only aircraft emissions which may have significant impact on air quality in the vicinity of Air Force bases.

Oxides of nitrogen are more likely to be emitted in the high power operating modes, such as the "military" mode. Emissions are typically between 20 and 30 lb/1000 lb of fuel, but, for some engines, may be as high as 300 lb/hr (Ref 1, 3, 5-9) reported as NO_2. Afterburners produce little additional nitrogen oxides. In general, the ratio of $NO:NO_2$ produced is about 10:1. The NO_x emissions are inconsequential in the idle mode.

Particulate emissions also are more likely to be produced in the higher power modes, and the amounts vary greatly among the various engine models. Afterburners consume some of the engine particulate matter, so engines operating in the afterburner mode have much lower particulate emissions (Ref 10).

It has been said that particulate emissions amount to about 0.5 percent of the fuel used (Ref 1). Actually, the value is probably a little high, being perhaps more nearly an upper limit. However, there are few reliable particulate measurements, although it is possible to approximate particulate emissions (Ref 11) from smoke number (Ref 12) data. The maximum particulate emission rate is about 50 lb/hr. Particulate emissions are mostly carbon (Ref 10), with particle sizes ranging from 0.01 to 1 micron. For test cells, particulate emissions are only important with respect to plume visibility, since there is only negligible impact on air quality (Ref 2).

Sulfur oxides arise from the sulfur in the fuel, which is about 0.06 percent by weight. In most instances, the preponderance of the sulfur oxides emitted will be SO_2, with only a small fraction as SO_3. However, more-or-less the opposite has been reported for an F100 engine test, where the ratio of $SO_3:SO_2$ was about 9:1 (Ref 3).

The sulfur oxide emissions from turbine engines are not considered to be a significant problem, especially with respect to the test cell emissions (Ref 2).

At the present, it appears that the plume opacity, or visibility, is the only area wherein turbine engine test cells have been cited for legal violations, although there is the possibility that NO_x emissions may sometimes exceed standards (Ref 1). Test cell emissions have little impact on air quality (Ref 2), so therefore it is the plume visibility which is of primary importance.

The Nature and Rationale for Test Cell Operations

The main purpose of the test cell is to enclose the engine under test so as to attenuate the noise. The usual test cell is a U-shaped building constructed mostly of concrete. The engine under test is mounted in the horizontal portion of the building, and its exhaust enters an augmentor tube. Air intake to the cell is through one of the vertical sections of the building, and the discharge stack is the other vertical section.

An important variable in the test cell operation is the operation of the augmentor, which is essentially an ejector energized by the high velocity exhaust. The augmentor serves to reduce the pressure behind the engine and thus to more nearly duplicate the conditions of the open air. Ideally, the augmentor reduces the pressure behind the engine to the same value as the pressure at the inlet. A variable amount of air is drawn into the augmentor along with the exhaust. In some cases, water is sprayed into the augmentor throat area to cool the exhaust and prevent damage to the structure from high temperatures. The major aim of this work is to explore the effects of the water addition on the cell operation and cell emissions.

The Test Cell Processes

The processes which occur in the test cell and which are of interest in this work include those occurring in the augmentor throat as gas and water make contact, those which occur in the augmentor diffuser section and those which occur in the cell discharge stack. These processes are all considered in order.

The process of predominate interest in the augmentor throat is the evaporation of the water droplets. Consideration is given to the evaporation rate which might be deduced from a prediction of the probable droplet size and probable flow conditions. An energy balance is used to predict the amount of water which evaporates and the resulting temperature which will prevail in the augmentor diffuser.

The processes of interest in the diffuser are condensation and particle wetting, droplet growth (with coagulation) and gas absorption in the droplet liquid. Consideration is given to all these processes and the results of each are predicted either by a model or by speculation based upon published work.

In the stack discharge, the process of additional gas absorption is considered, as well as deposition of the droplets on the wall surfaces of the stack and baffling. Relations to predict the final emissions of all pollutants in the gas phase and in the liquid phase

are developed.

Finally, methods of predicting plume visibility are adapted from published works and applied to the cell emissions.

2. DESCRIPTION OF THE PROCESS

Augmentor Mixing and Evaporation Rates

As the exhaust exits the engine, its kinetic energy carries it into the throat of the augmentor. The resulting pressure reduction causes additional dilution air to also be directed into the augmentor. The dilution air flow varies from 0.5 to 2.5 times the exhaust rate (mass basis). Presumably the jet rapidly dissipates by mixing with the dilution air.

An exhaust discharged horizontally to the open atmosphere will maintain some unmixed potential flow for about six jet diameters downstream (Ref 10). This typically corresponds to two to four jet diameters after entering the augmentor. Mixing with dilution air would normally be very rapid, especially if a water spray is introduced almost immediately inside the augmentor throat.

The velocity of gases (exhaust plus dilution air) are typically about 500 ft/sec in the throat at the point of the first water spray contact. For such a situation, it is possible to estimate the maximum droplet size of the water from the Weber Number. According to Perry (Ref 13) the maximum stable droplet size may be predicted by assuming a Weber Number, We, of 10 for short duration shock contact with high velocity gas.

The Weber Number is given by:

$$We = \frac{U_g^2 \rho_g D\rho}{\delta g_c}$$

If a Weber Number of 10 describes this situation, the maximum droplet diameter can be predicted by:

$$D_{max} = \frac{10\ \delta\ g_c}{U_g^2\ \rho_g} \qquad (1)$$

A typical case (U_g = 500 ft/sec, ρ_g = 0.04 lb/ft^3) would suggest water droplet diameters of about 50 micron, and less.

As water droplets are further broken up by the gas stream, a volume mean diameter can be predicted (Ref 13) as:

$$D_{vm} = \left(\frac{136\ U_e\ \rho^{3/2} D_I^{1/2} g_c^{3/2}}{\rho_g^2\ \rho_w^{1/2}\ U^4} \right)^{1/3}, \text{ where} \qquad (2)$$

D_{vm} is the volume mean diameter of the droplet population.

Assuming that $D_I \sim D_{max}$, the volume (also mass) mean diameter for the case suggested above is about 14 micron, which provides a surface area of approximately 2000 ft^2/lb.

The droplet Reynolds Number will be about 75, which would suggest a heat transfer coefficient of about 0.7 BTU/sec.ft^2.^oF and an initial evaporation rate of about 550 lb/sec per lb of liquid water. Thus, at the initial rate, a droplet would evaporate in less than 0.002 seconds, or by the time it had travelled about a foot at the example velocity of 500 ft/sec. It seems appropriate, therefore, to assume the water is either all evaporated, or the excess liquid and gas are in equilibrium as the stream exits the throat of the augmentor.

Energy Balance

The relationship between the amount of water which evaporates and the resulting temperature can be determined from an energy balance. The maximum amount of water evaporation possible may be determined from considerations of psychrometric principles.

The energy balance around the augmentor can be simplified by the following initial assumptions: the water and air both enter at the same ambient temperature, T_o; the enthalpies of both air and liquid water are taken as zero at T_o; the heat capacities are treated as constant; and friction, as well as changes in elevation and pressure, are ignored.

For the case of the turbojet engine (or a turbofan engine if the bypass is included with the exhaust), no shaft work crosses the system boundary. Therefore, the enthalpy and kinetic energy of the total exhaust is equal to the heating value of the fuel, the process being nearly adiabatic. Thus: a general energy balance statement is:

$$FQ_F = H_g\ m_g + H_v\ m_v + m_w\ H_w + \frac{(m_g + m_v + m_w)U^2}{2g_c} \qquad (3)$$

$m_g = m_e + m_a$, and $m_a = m_e R_d$,

so $m_g = m_e (R_d + 1)$,

Further, $m_e = F (1 + A/F) = F \left[1 + 1/(F/A)\right]$

Thus: $m_g = F \left[1 + 1/(F/A)\right] (R_d + 1)$

(Note: Ref 11 presents a method for determining m_g from the fuel rate and cell depression pressure, when cell geometry is known. This information may sometimes be more readily available than R_d.)

If this system has attained equilibrium at temperature T,

$H_g = Cp_g (T-To)$

$H_v = \lambda + Cp_v (T-To)$

$H_w = Cp_w (T-To)$

Substituting in the general energy balance statement:

$$FQ_F = (T-To)(Cp_g m_g + Cp_v m_v + Cp_w m_w) + \lambda m_v + (m_g + m_v + m_w)U^2/2g_c \quad (4)$$

At the equilibrium condition, the temperature will be approximately the adiabatic saturation temperature, the small difference being due to the sensible heating of the excess liquid water from ambient to the mixture temperature.

The average velocity can be deduced by establishing an arbitrary finite time, such as one second. Then, F, m_v, and m_w are taken as mass flow rates, for example, kg/sec.

The total gas volume is:

$$V_g = \frac{m_g RT}{P M_g} \quad (5)$$

The vapor volume:

$$V_v = \frac{m_v RT}{P M_w} \quad (6)$$

The velocity can be expressed as:

$$U = \frac{(V_g + V_v) \times 4}{\pi D_t^2} \tag{7}$$

This expression ignores the volume of the excess liquid.

The adiabatic saturation temperature (T_s) depends upon the inlet (to the augmentor) condition of the gas, i.e., the temperature and water vapor content.

The water vapor content of the entering gas can be predicted from the ambient humidity, the air and fuel rates and the H/C ratio of the fuel, thus:

The moles of air supplied to the engine and the augmentor are:

moles air = $\frac{m_g - F}{29}$ where the molecular wt of the air is taken as 29.

If the ambient molal humidity (moles water/mole air) is h, then, the moles of water supplied with the air are:

moles water = $h\,(m_g - F)/29$

The water of combustion, for fuel with a H/C (atomic) ratio of α, is:

moles water = $\frac{1}{2}\left(\frac{\alpha}{12 + \alpha}\right)F$

Total vapor phase water entering the augmentor:

$$\text{moles water} = \frac{h\,(m_g - F)}{29} + \frac{1}{2}\left(\frac{\alpha}{12 + \alpha}\right)F$$

The water evaporated has previously been described as m_v (mass) so the total water vapor, (moles) in the augmentor is:

$$W = \frac{h\,(mg - F}{29} + \frac{1}{2}\left(\frac{\alpha}{12 + \alpha}\right)F + \frac{m_v}{18} \tag{8}$$

The gas in the augmentor has previously been described as (mass):

$$m_g = m_e (1 + R_d)$$

The value of m_e as previously defined does include the water of combustion, but not the humidity of the ambient air. If the molecular weight of the exhaust and air are both taken as 29, then the molar amount of <u>dry</u> augmentor gas is:

$$\text{moles dry gas} = \frac{m_g}{29} - \frac{1}{2}\left(\frac{\alpha}{12 + \alpha}\right) F$$

The mole fraction of water in the augmentor is

$$Y = \frac{W}{\frac{m_g}{29} - \frac{1}{2}\left(\frac{\alpha}{12 + \alpha}\right) F} \qquad (9)$$

The partial pressure of the water vapor is:

$P_v = P_t y$, where P_t is the pressure in the throat.

The adiabatic saturation equilibrium will occur as the quantity of water evaporated reduces the temperature of the gas while also increasing the water vapor content until the partial pressure of the water vapor, P_v, is equal to the vapor pressure of water, p^*.

In the range or temperatures of interest in this study (30 to 60° C), the vapor pressure of water can be approximated by:

$p^* = \exp(4.3320 - 843.084/T)$ with p^* in kPa and T in °K. (This relation is satisfactory only for the limited temperature range suggested above.

The saturation temperature may often be higher than the ambient, or water, temperature, hence the gas stream can possibly be cooled below the adiabatic saturation temperature (T_s). The effect is expected to be small because of the thermal effects of the condensation. There is no convenient single relation to compute the adiabatic saturation temperature, T_s. However, the mole fraction of water vapor calculated by eqn (9) will be:

$$y_{Ts} = \frac{p^*_{Ts}}{P} \tag{9a}$$

A value of T_s which satisfies eqns (4), (9) and (9a) may be found by iterative procedures.

The pollutant loadings in the gas stream can be determined from the engine emission factors and the amount of fuel used. Emission factors relate the mass of pollutant per mass of fuel, usually in lb/1000 lb of fuel. Thus, for pollutant component j,

$$m_j = \frac{E_j F}{1000} \tag{10}$$

For particulate matter, the emissions may be approximated from the method of Finch and Eyl (Ref 11). Their expression,

$$\frac{\omega_e}{\rho_e} \cdot \frac{W}{A} = 10^{(0.0263/SN) - 5.96)} \tag{11}$$

is valid for smoke number values (SN) greater than 15. The term $\frac{W}{A}$ is a sample collection parameter, mass of sample per unit area of filter as specified by the smoke number method (Ref 12). It should be noted that $\frac{\omega_e}{\rho_e}$ is the mass ratio of particulate matter to exhaust gas.

The total particulate emission may be estimated from

$m_p = \frac{\omega_e}{\rho_e} m_e$, with m_e estimated from the fuel and fuel/air ratio.

The particulate emission factor is, then,

$$E_p = 1000 \frac{\omega_e}{\rho_e} \frac{m_e}{F} \tag{12}$$

Discharge Into Diffuser

As the gas-vapor stream discharges from the throat of the augmentor into the diffuser section, a velocity change will occur as a consequence of the change in cross-sectional area. If the flow in the throat is subsonic, as is likely, then the increased cross sectional area of the diffuser will cause a pressure increase as predicted by the Bernoulli theorem.

The increase in pressure brought about by the velocity change causes condensation of some water vapor. The condensation will simultaneously cause an increase in temperature as the latent heat of the vapor is given up. The energy balance of eqn (4) and eqns (9) and (9a) apply equally well to the diffuser section. The computations are more appropriate, in fact, in the diffuser section, since the temperature and water vapor content will change only very little from the diffuser to the ultimate discharge.

Flow Through The Diffuser

Diffuser velocities are relatively high, probably of the order of 200 ft/sec ($\sim$ 60 m/s). It is expected that some minor interaction between the particles, or droplets, and the walls occurs. Whereas droplets will impinge on the wall, the velocity may be great enough to re-entrain most of the water which is collected. The re-entrainment aspects are not at all well known. Actually, the rate of deposition of the droplets on the wall is not readily predictable with confidence, although the problem has been considered in the literature.

Any attempt at evaluation of the effect of liquid water droplets in the diffuser must begin with a consideration of the probable droplet diameter. Robsen, et al (Ref 1) have suggested that condensation in the test cell environment will be by nucleation on the carbon particles, and not by condensation on the hygroscopic gases (SO_3 and NO_2). The same reference reports that fog droplets tend to have a mean diameter of 38 microns with a standard deviation of ± 12 microns. The ultimate diameter of 38 microns, however, does not necessarily occur immediately, and the initial droplet formation probably will result in droplets of 0.1 to 1.0 microns surrounding the carbon (smoke) particles which serve as nuclei.

The amount of water condensed is ordinarily less than one-tenth of one percent of the gas rate (mass). The excess liquid, however, may be as much as ten to fifteen percent of the dry gas rate, although it is of course highly variable, depending upon the operation. Probably, though, in any situation where appreciable liquid is present, the droplet size is primarily a function of the droplet size of the excess liquid, since it will be most of the liquid present.

The coagulation of the smaller particles in the highly turbulent state can only be guessed. Some work has been done, much of it reported by Fuchs (Ref 14), but there is little which can be used to relate to conditions different from the experimental.

If the only liquid present were from condensation, then coagulation would not be a significant process. In such a case, one could expect a droplet diameter increase of only about ten percent at most during passage through the diffuser.

On the other hand, if as was suggested, the excess spray is the predominate source of liquid, a mean diameter of 20 microns seems to be a reasonable, but speculative, estimate of the droplet diameter in the diffuser. It should be noted that a droplet size of about 14 micron was predicted for the formation of the droplets in the throat.

Absorption Of Gases In The Diffuser

The liquid water droplets, either formed from condensation or present as unevaporated excess will absorb gases in the diffuser during passage through it. In some circumstances this might be an important process either in relation to the amount of a gas removed from the gas stream or because of the amount of acid forming solute added to the liquid phase.

The gases which might be absorbed include CO_2, CO, NO, NO_2, SO_2, SO_3, and HC, the first three listed having very low wolubility. The effect of the water on the CO_2 content would be expected to be negligible. CO_2 is not considered to be a pollutant in this study.

It can be easily shown by example that the effect of water on the NO emissions is also negligible.

An example can be taken from Ref 3. In this case, an F100 engine was emitting about 36.6 lb/hr of NO, there were approximately 400 gallons per minute of water being discharged as a mist in a gas stream (mostly air) of 2,289,000 lb/hr. The temperature was approximately 100°F.

From these data, it may be deduced that the mole fraction of NO in the exiting gas-stream was 1.5×10^{-4}. The Henry's Law constant for NO at about this temperature (40°C) is given by Perry (Ref 15) as 3.52×10^{4} atm/mole fraction. Therefore if 400 gallons per minute of water ($\sim$ 11,000 lb mole/hr) came to equilibrium with the initial NO concentration, it would dissolve only 0.0014 lb of NO, which is less than 0.005 percent of the total NO emitted. Similar order of magnitude results will be found for the case of CO, and therefore, it is concluded that the absorption of neither NO nor CO is of consequence in this study.

A calculation similar to the above for the case of SO_2 reveals

that about 0.28 percent of the SO_2 would be dissolved in the water if equilibrium were achieved. This too, seems to be unimportant, and it is concluded that the water spray has only negligible effect on the SO_2 emissions.

The hydrocarbons present special problems of prediction. Some of the unburned hydrocarbons are quite soluble, while others are not. Conkle, et. al. (Ref 17), report an analysis of the exhaust hydrocarbons from a T-56 combustor. From these results it appears that about 25 percent of the HC could be considered as soluble, with acetaldehyde being a prominent and perhaps typical soluble constituent.

Presumably about 75 percent of the unburned hydrocarbons are unaffected by the water spray.

For the cases of SO_3, NO_2, and the soluble hydrocarbons, it is not appropriate to treat the solubilities in the same manner as above since these pollutants are highly soluble in water and would not obey Henry's Law. For these gases, it may be estimated that the Henry's Law constant is approximately zero, so they will not come to "equilibrium" in the usual sense, at least not in this dilute situation. The limiting amount of gas dissolved in the mist droplets will be a function of the rate of mass transfer.

Mass transfer rate predictions for droplets travelling in a gas stream have been developed (Ref 18) for the case of turbulent flow around the drops. Since in the present case the mist droplets are moving at essentially the same velocity as the gas stream, the principle transport mechanism is probably simply molecular diffusion. The turbulence in the system might increase the mass transfer rate, even though the overall relative velocity of the gas with respect to the mist droplets is zero. If the mechanism of transport is simply molecular diffusion, the diffusion rate can be written for this extremely dilute case as:

$$N_A = D_v \frac{dc}{dx} \tag{13}$$

where $\frac{dc}{dx}$ is the concentration gradient that exists between the liquid surface and the bulk of the gas.

The concentration gradient can be evaluated by making the assumptions that the diffusion path length is one-half the distance between adjacent droplet surfaces and that the interface composition is zero (essentially infinite solubility and complete mixing in the droplet).

Concentration is in moles/unit volume, hence, for component j:

$$c_j = \frac{m_j}{M_j \; V_g} \tag{14}$$

The volume of the gas, V_g, is as given by eqn (5).

The bulk phase concentration is, then

$$c_j = \frac{m_j \; P \; M_g}{M_j \; m_g \; RT} \tag{15}$$

The mean spacing between adjacent droplets can be approximated from the volume fraction of droplets and the droplet diameter. The ratio of void (gas) to non void will be approximately the cube root of the ratio of void volume to droplet volume, this latter ratio being:

$$\text{Volume ratio} = \frac{m_g \; RT \; \rho_w}{M_g \; P \; m_w} \tag{16}$$

Thus, the ratio of spacing to droplet diameter is

$$\frac{\text{spacing}}{\text{diam}} = \left(\frac{m_g \; RT \; \rho_w}{M_g \; P \; m_w} \right)^{1/3} \tag{17}$$

and for droplet diameters, D_p, the diffusion path length is:

$$x = 1/2 \left(\frac{m_g \; RT \; \rho_w}{M_g \; P \; m_w} \right)^{1/3} D_p \tag{18}$$

The diffusion equation becomes:

$$(N_A)_j = (D_v)_j \; x \; \frac{2 \; m_j \; PM_g}{M_j \; m_g \; R \; T \; D_p} \left(\frac{M_g \; P \; m_w}{m_g \; RT \; \rho_w} \right)^{1/3} \tag{19}$$

$(N_A)_j$ is the absorption rate per unit area, and is the rate of disappearance of component j from the gas phase, therefore:

$$(N_A)_j = - \frac{1}{M_j A} \frac{d\, m_j}{dt}$$

The area, A is the total droplet surface area, and can be given by:

$$A = \frac{6}{D_p \rho_w} m_w$$

Combining terms and separating variables

$$\frac{d\, m_j}{m_j} = - 12 (D_v)_j \frac{P M_g}{m_g RT D_p^2 \rho_w} \frac{m_w}{} \left(\frac{m_g P m_w}{M_g RT \rho_w} \right)^{1/3} dt \qquad (20)$$

Integrating this relation over the time required to pass through the diffuser gives:

$$\ln \left[(M_j)_2 / (M_j)_1 \right] = \frac{-12 (D_v)_j P M_g m_w}{m_g RT D_p^2 \rho_w} \left(\frac{m_g P m_w}{M_g RT \rho_w} \right)^{1/3} t \qquad (21)$$

Where t = time to pass through the diffuser (L_d/U_d), $(m_j)_1$, $(m_j)_2$ = mass of component j in gas, initially and after time t. The diffusivities of the gases of interest, NO_2, SO_3 and acetaldehyde, as representative of soluble hydrocarbons in air can be estimated by the method of Gilliland (Ref 16). At 40°C, the values are:

for NO_2, $D_v = 0.134$ cm^2/sec

for SO_3, $D_v = 0.099$ cm^2/sec

for acetaldehyde, $D_v = 0.14$ cm^2/sec

The temperature of 40°C is probably close to the process temperature in most cases, but the diffusivities vary approximately as the 3/2 power of the absolute temperature, so may be estimated for any temperature of interest.

Discharge From The Augmentor

Discharge from the augmentor is accompanied by considerable turbulence and involves flow through perforated walls or other types of baffling, and the accompanying agitation and mixing is assumed to result in a sharp reduction in velocity with not much change in pressure. The increased turbulence will cause some extra coagulation, and it might be assumed that a stable dispersion of droplet diameter of ∿38 microns (Ref 1) will be formed.

Flow Through The Stack

Stack velocities are relatively low, on the order of 5 to 15 m/s (about 15 to 50 ft/sec). The processes which occur are very similar to those which occur in the diffuser section of the augmentor. The stack velocity will ordinarily be in excess of the settling velocity of any droplet or particle, so no settling will occur.

The absorption of the soluble gases will continue if this process is not essentially completed in the diffuser. The amount of any soluble component absorbed may be predicted by the use of eqn (21), except for a change in subscript on mj, such that:

$$\ln\left[\frac{(m_j)_3}{(m_j)_2}\right] = -\frac{12(D_v)_j \, PM_g \, m_w}{m_g \, RT \, D_\rho^{\,2} \rho_w} \left(\frac{m_g P \, m_w}{M_g \, RT \, \rho_w}\right)^{1/3} t \qquad (21a)$$

where now t is the time for passage through the stack and $(m_j)_3$ is the final mass of gas phase component j.

Liquid Deposition in the Stack

As the gas-mist mixture flows out the stack, some of the particles will strike the wall and remain, forming a liquid film, eventually to flow downward by gravity or upward because of drag from the main stream. In most cases the result of this deposition, in terms of the fraction of liquid retained, will be minimal, perhaps negligible. However, for the possible occasional case where the effect is of interest, it is possible to estimate the effect. The rate of particle (or droplet) deposition on the walls of the stack is difficult to evaluate with confidence, especially if the droplet diameter is questionable. However, Fuchs (Ref 14) does present a relation for approximating the rate of impaction of particles on walls during turbulent flow. The actual measured rates have been observed to be from one-half to two times the prediction. The relation given by Fuchs is:

$$I = \frac{N U^{*2}}{U\left[1 + U^{*}\left(14.5^{3}/2x^{*3} - 50,6\right)\right]} \qquad (22)$$

I = Impacting flux (particles/unit area and time)

N = Particle population, (particles/unit volume)

U = Velocity of the stream

U^* = Eddy velocity

$$U^{*} = \frac{0.2\,U}{(Re)^{1/8}}$$

x^* = $x_i\, U^*/\nu$ (dimensionless stopping distance)

x_i = Stopping distance

ν = Kinematic viscosity

$$x_i = \frac{2\,U_I\, r_p^{2}\, \rho_p}{9\mu}$$

U_I = Particle initial velocity

Presumably $U_I \sim U^*$.

The particle population density, N, is the total number of droplets divided by the total volume, or the droplet flow rate divided by the volumetric flow rate.

The total droplet number flow rate is the liquid mass rate divided by the mass of a single droplet. The mass of a droplet is:

$$mp = \frac{\pi D_p^{3} \rho_p}{6} \qquad (23)$$

The liquid mass rate is m_w, so the particle number flow rate is $6\, m_w/\pi D_p^{3}\rho_p$.

Thus, since the total volume is Vg + Vv, the particle population density is:

$$N = \frac{6\, m_w}{\pi (V_g + V_v)\, D_p^{\,3} \rho_p} \qquad (24)$$

The rate of mass removal is the rate of droplet removal multiplied by the mass of a single droplet if it is assumed that reentrainment is negligible. The rate of droplet removal is the impacting flux, I, multiplied by the wall surface area, A_s. Thus:

$$\frac{dm_w}{dt} = -I\, A_s \times \frac{\pi\, D_p^{\,3}\, \rho_p}{6} \qquad (25)$$

Therefore:

$$\frac{dm_w}{dt} = -\frac{As\, \pi\, D_p^{\,3} \rho_p}{6} \times \frac{6\, m_w}{\pi (V_g + V_v) D_p^{\,3} \rho_p} \times \frac{U^{*2}}{U\left[1 + U^*(14.5^3/2x^{*3} - 50.6)\right]}$$

Upon simplification and separation of variables:

$$\frac{dm_w}{m_w} = -\frac{As\, U^{*2}\, dt}{(V_g + V_v) U\left[1 + U^*(14.5^3/2x^{*3} - 50.6)\right]} \qquad (26)$$

Integrating over the time in the stack, t, gives:

$$\ln\left(\frac{Mw,o}{M_w}\right) = -\frac{A_s\, U^{*2}\, dt}{(V_g + V_v) U\left[1 + U^*(14.5^3/2x^{*3} - 50.6)\right]} \qquad (27)$$

Where $M_{w,o}$ is the mass rate of water droplets actually discharged from the stack. It is to be noted that this ignores the amount deposited in the diffuser which is assumed to be not only a negligible amount, but probably mostly re-entrained as well.

The fraction of particulates deposited with the water is assumed to be proportional to the fraction of water deposited, thus the particulate mass emitted is;

$$mp,o = M_p \left[1 - \frac{M_{w,o}}{m_w}\right] \tag{28}$$

where $m_p = \frac{Ep\ F}{1000}$ as in eqn (10), and Ep is as given by eqn (12).

The mass of any soluble gas emitted in the gas phase is $(mj)_3$ as given by eqn (21a). The amount of soluble gas absorbed in the stack is given by

$$(mj)_{w,s} = (mj)_2 - (mj)_3.$$

The mean amount of soluble component j dissolved in the liquid during passage through the stack is approximated by

$$\overline{(mj)}_{w,s} = \frac{mj - (mj)_2 + mj - (mj)_3}{2} = \frac{2\ mj - (mj)_2 - (mj)_3}{2} \tag{29}$$

The amount of soluble material j deposited in the stack is:

$$(mj)_{d,s} = (\overline{mj})_s \times (m_w - m_{w,o}) \tag{30}$$

The amount of soluble material emitted in the mist is given by:

$$(mj)_{w,o} = mj - (mj)_3 - (mj)_{d,s} \tag{31}$$

3. PLUME VISIBILITY

The visibility of the plume may be considered on the basis of either a white plume or a dark plume. If the exhausted gases are saturated with water vapor at 40° to 50° C, the plume will become highly visible because of the condensation of vapor caused by sudden contact with cooler outside air. The plume will appear white because of the scattering of light reflected from the droplets of newly condensed water. This steam plume will usually obscure the other components of the plume.

The presence of a steam plume is almost always exempted from legal considerations. Nevertheless, for the possible occasional interest in steam plume visibility, an approximation of the Ringelmann Number may be made.

The general problem of predicting white plume optical properties is very complex. The topic has been considered by Halow and Zeek (Ref 19) who developed a Ringelmann Number correlation. The problem is complicated by variables of weather, besides the complication of estimating the droplet size.

The steam plume results from condensation in the cool air, and for the purposes of this study, a "typical" condition is assumed for comparison. These typical conditions are a clear day, with viewing at a scattering angle of 120°, which corresponds to a common Ringelmann reading condition (Ref 18). Other assumptions:

Fog droplet diameter ~ 2.0 micron with a standard deviation of ± 1.0 micron (typical from Ref 17).

All "excess" vapor condenses, i.e., the stack gases become saturated at Ta (Ta is the common ambient temperature). None of the newly formed droplets re-evaporate.

These conditions correspond to approximately a "worst" case, except for the viewing angle.

Ref 18 presents generalized curves, developed from computations based on the Mie theory, to predict a ratio of scattered brightness (Bs) to ambient brightness (Bo) for a unit particle loading, and a unit stack diameter. For the conditions proposed here, the ratio is 5.5. The value 5.5 is employed with the actual particle loading and actual stack diameter as follows:

$$Bs/Bo = 5.5 \times 10^{7} D_s' \omega' / \rho_p' \qquad (32)$$

This is a dimensional equation, where

D'_s = stack diameter, ft

ω' = particle loading, lf/ft^3

ρ'_p = particle density, lb/ft^3

The value of (Bs/Bo) so obtained is then used to predict the Ringelmann Number. As might be expected, a great deal of scatter is apparent in the Ringelmann Number versus (Bs/Bo) correlation. However, the best fit curve presented by the authors can be approximated by:

$$R_n = 0.0439\ (BS/Bo) - 8.64 \times 10^{-5}\ (Bs/Bo)^2 \qquad (33)$$

Eqn (33) predicts the Ringelmann number best fit curve up to a value of (Bs/Bo) = 140, where $R_n \sim 4.5$, which is about a maximum white plume Ringelmann Number.

In the use of eqns (32) and (33), it is probably appropriate to use the largest dimension of a rectangular stack for D'_s, and to use only the newly condensed water for ω', since the larger droplets leaving the stack will have little effect on the white plume visibility.

Since a steam plume is not usually objectionable, even though highly visible, it is probably more realistic to estimate plume visibility as that which may be deduced for the spray water discharge as a "black" plume.

Ref 17 also presents a best fit correlation for black plume Ringelmann numbers, as a function of transmittance. The curve can be approximated by three different equations depending upon the transmittance, as follows:

$$Rn = 5 - 4.8\ Tr \text{ for } 0 < Tr < 20\% \qquad (34a)$$

$$Rn = 4.5 - 0.037Tr \text{ for } 20 < Tr < 85\% \qquad (34b)$$

$$Rn = 8.8 - 0.088Tr \text{ for } Tr < 85\% \qquad (34c)$$

Eqns (34a, b, c) may be used for comparison with dry, or carbon plumes, as for example, with the transmittance predictions of Finch

and Eyl (Ref 11), and might also be used for the wet plume in some circumstances.

The transmittance for the dark plume case may be approximated from the Beer-Lambert Law, thus;

$$Tr = \exp - (S_p \, \omega_{w,o} D_s) \tag{35}$$

The value of S_p, of course, is difficult to evaluate accurately. A value of 400 cm^2/g, based on 38 micron spheres, is probably a conservatively high estimate, since the extinction efficiency is probably less than unity. On the other hand, the diameters might be less than 38 microns.

The value of $\omega_{w,o}$ can be deduced from

$$\omega_{w,o} = \frac{m_{w,o}}{V_g} \tag{36}$$

For both the black and the white plume calculations, it is appropriate to use the largest stack dimension for the stack diameter.

4. SAMPLE CALCULATIONS

A. INPUT DATA*

Engine:	F100-PW100
Fuel Rate:	10300 lb/hr; H/C = 2.02
Fuel Heating Value:	19,000 Btu/lb
Fuel/air Ratio;	0.013
Dilution Ratio (Augmentor)	1.85
Emission Factors:	(lb/1000 lb of fuel)
	HC; 0.30; CO: 9.5; NO: 231.7;
	NO_2: 45.1; SO_2: 0.1; SO_3: 1.1
Smoke No = 45.1 @ W/A = 0.023	lb/in^2
Cooling Water Flow	300,000 lb/hr (600 gpm)

Augmentor Dimensions

Throat diam:	6 ft
Diffuser diam:	8 ft
Diffuser length	30 ft

Discharge Stack Dimensions:

12 parallel channels, each 17.5 inches wide by 13.5 ft long and 13.0 ft deep. (Total wall area $\sim$ 236 ft^2).

Ambient temperature	80°F (26.7°C)
Barometric Pressure	1 std atm
Relative humidity	50 percent

*This data set is approximately as given for "Run 127" of Ref 3.

B. COMPUTE SATURATION TEMPERATURE

(Compute diffuser temperature, since diffuser pressure is assumed to be barometric pressure).

h = 0.017 mole water vapor/mole air

m_g = 288.2 kg/s; F = 1.30 kg/s

Cp_g = 6.95 cal/g mole °K = 1003 j/kg °K

Cp_v = 0,44 cal/g °K = 1842 j/kg °K

Cp_w = 1,0 cal/g °K = 4187 j/kg °K

λ_{H2O} = 1047 Btu/lb = 2,435 x 10^6 j/kg

FQ_F = 1,957 x 10^8 Btu/hr = 5,74 x 10^7 j/s

Simultaneously satisfy eqns (4), (8), and (9) by iterative calculations

$$W = \frac{h(m_g - F)}{29} + \frac{1}{2}\,\frac{\alpha}{12+\alpha}\,F + \frac{m_v}{18} = 0,262 + \frac{m_v}{18}$$

V_g = 0,816 x T m^3/s (T in °K)

V_v = 0.0046 x m_vT m^3/s (T in °K, m_v in kg/s)

M_w = 37,8 - m_v kg/s

Assume T = 40°C

Vg = 0,816x(40+273) = 255,4 m^3/s

U_d ≃ 54.7 m/s (ignores water volume)

5.74 x 10^7 = (40-26,7) 1003 x 288.2 + 1842 m_v + 4187x(37.8-m_v) + 2,435 x 10^6 m_v + (288,2 + 37.8) x $(54.7)^2/2$

m_v = 21.2 kg/s

W = 1.44 k mole/s

y = 0.1255

P* (a + 40°C) = 7.20 k P_a

$P^*/p = \frac{7.20}{101,3} = 0.071$

Therefore, since P*/p < y, 21.2 kg/s cannot be evaporated, so Ts (and therefore, T) will be greater than 40°C.

By trial, $T_s \simeq 50^{\circ}C$, Ud = 64 m/s , m_v = 19.3 kg/s, y = 0.119 mole fraction water vapor, m_w = 37.8, m_v = 18.5 kg/s

C. COMPUTE GAS ABSORPTION IN DIFFUSER

$$\ln \frac{(M_j)_2}{(m_j)_1} = -12(D_v)_j \frac{PM_g \; m_w}{mg \; RT \; D_p^2 \; \rho w} \left(\frac{M_g \; P \; m_w}{m_g \; RT \; \rho w} \right)^{\frac{1}{3}} t$$

$M_g \simeq 29$

$\rho_w = 990 \text{ kg/m}^3$

$D_p = 2 \times 10^{-5}$ m

$(D_v)_{NO2} = 0.134 \text{ cm}^2/\text{s} = 1.34 \times 10^{-5} \text{ m}^2/\text{s}$

$(D_v)_{SO3} = 9.9 \times 10^{-6} \text{ m}^2/\text{s}$

$(D_v)_{Acet} = 1.4 \times 10^{-5} \text{ m}^2/\text{s}$

$(m_{NO2})_1 = 0.0585$ kg/s

$(m_{SO_3})_1 = 0.00143$ kg/s

$(m_{Acet})_1 = 0.000097$ kg/s

t = 0.14s

$$\left(\frac{m_2}{m_1} \right)_{NO_2} = 0.85$$

$$\left(\frac{m_2}{m_1} \right)_{SO_3} = 0.89$$

$$\left(\frac{m_2}{m_1} \right)_{\text{Acetaldehyde}} = 0.84$$

AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

2 OF 6
AD A061624


624

D. Compute Gas Absorption In Stack

$D_p = 3.8 \times 10^{-5}$ m

$U_s = \frac{V_g + V_v}{S_s} = \frac{292.2 \text{ m}^3/\text{s}}{22 \text{ m}^2} = 13.3$ m/s

$t_s = 0.286$ s 0.30

$\left(\frac{m_3}{m_2}\right)_{NO_2} = 0.91$

$\left(\frac{m_3}{m_2}\right)_{SO_3} = 0.93$

$\left(\frac{m_3}{m_2}\right)_{Acel} = 0.91$

Fraction of soluble gases remaining in vapor phase:

NO_2: 0.85 x 0.91 = 0.77

SO_3: 0.89 x 0.93 = 0.83

HC : 0.84 x 0.91 = 0.76 (soluble)

Total HC, assuming 75 percent insoluble: 0.94

E. Compute Water Deposit In Stack

Equiv diam = $4R_h = 4 \times \frac{0.445 \times 4.116}{2(0.445 + 4.116)} = 0.803$ m

$\mu \sim 1.8 \times 10^{-5}$ kg/ms

$\rho = 1.05$ kg/m^3

$U = 13.3$ m/s

$(V_v + V_g) = 292.2$ m^3/s

$Re_s = 6.23 \times 10^5$

$$U^* = \frac{0.2 \times 13.3}{(6.23 \times 10^5)^{\frac{1}{8}}} = 0.5 \text{ m/s}$$

$$x^* = 64.35$$

$$A_{w,s} = 43.4 \text{ m}^3$$

$$\ln \frac{m_{w,o}}{m_w} = - \frac{U^{*2} A_{w,s} t}{(V_g + V_v) U\left[1 + U^*(14.5^3/2x^{*3} - 50.6)\right]}$$

$$\frac{M_{w,o}}{m_w} = 0.99$$

(0.01 mass fraction (1 percent) deposited.)

6. Compute Gas Phase Emissions

In gas phase:

$$(m_j)_{g,o} = m_j \times \frac{(m_j)_2}{(m_j)_1} \times \frac{(m_j)_3}{(m_j)_2}$$

$$(m_{NO2})_{g,o} = 0.0585 \times 0.85 \times 0.91 = 0.045 \text{ kg/s}$$

$$(m_{SO_3})_{g,o} = 0.00143 \times 0.89 \times 0.93 = 0.0012 \text{ kg/s}$$

(soluble) $(m_{HC})_{g,o} = 9.8 \times 10^{-5} \times 0.84 \times 0.91 = 7.4 \times 10^{-5} \text{kg/s}$

(insoluble) $(m_{HC})_{g,0} = 2.9 \times 10^{-4} \text{ kg/s}$

$$(m_{CO})_{g,o} = m_{CO} = 0.0123 \text{ kg/s}$$
$$(m_{NO})_{g,o} = m_{NO} = 0.301 \text{ kg/s}$$
$$(m_{SO2})_{g,o} = m_{SO_2} = 1.3 \times 10^{-4} \text{ kg/s}$$

F. Compute Liquid Phase Emissions (Soluble Gases):

$$(m_j)_2 = m_j \times \frac{(m_j)_2}{(m_j)_1} \; ; \; (m_j)_3 = (m_j)_2 \times \frac{(m_j)_3}{(m_j)_2}$$

$(m_{NO_2})_2 = 0.05$ kg/s; $(m_{NO_2})_3 = 0.045$ kg/s

$(m_{SO_3})_2 = 1.3 \times 10^{-3}$ kg/s; $(m_{SO_3})_3 = 1.1 \times 10^{-3}$ kg/s

$(m_{HC})_2$(soluble) $= 8 \times 10^{-5}$ kg/s; $(m_{HC})_3 = 7.4 \times 10^{-5}$

Avg. amounts in solution during pass through stock:

$$(\overline{mj})_{w,s} = \frac{2\,mj - (mn)_2\,(mj)_3}{2}$$

$(\overline{m_{NO_2}})_{w,s} = 0.011$ kg/s

$(\overline{m_{SO_3}})_{w,s} = 2.3 \times 10^{-4}$ kg/s

$(\overline{m_{HC}})_{w,s} = 2 \times 10^{-5}$ kg/s

Soluble gases deposited in stack:

$$(m_j)_{d,s} = (\overline{m_j})_{w,s} \times 1 - \frac{m_{w,o}}{mw}$$

$(m_{NO_2})_{d,s} = 1 \times 10^{-4}$ kg/s

$(m_{SO_3})_{d,s} = 2.3 \times 10^{-6}$ kg/s

$(m_{HC})_{d,s} = 2 \times 10^{-7}$ kg/s

Soluble gases emitted in liquid phase:

$$(m_j)_{w,c} - m_j - (m_j)_{g,o} - (m_j)_{d,s}$$

$$(m_{NO_3})_{w,o} = 0.0134 \text{ kg/s}$$

$$(m_{SO_3})_{w,o} = 0.000228 \text{ kg/s}$$

$$(m_{HC})_{w,o} = 0.000024 \text{ kg/s}$$

G. Compute Particulate Emissions

$$\frac{\omega_e}{\rho_e} \cdot \frac{W}{A} = 10^{(0.0263(SN) - 5.96)}$$

$$\frac{\omega_e}{\rho_e} = 0.000732$$

$$Ep = 1000 \times 0.000732 \times 8.03 \times 10^5/10,3000$$

$$Ep = 57 \text{ lb/1000 lb Fuel}$$

$$Mp = 0.16 \text{ kg/s}$$

$$m_{p,o} = 0.16 \times 0.99 = 0.16 \text{ kg/s}$$

H. Compute Ringelmann No (Black Plume):

$$Tr = \exp - (Sp\ \omega_{w,o}\ D_s)$$

$$D_s = 8.23 \text{ m}$$

$$W_{w,o} = 0.0633 \text{ kg/m}^3$$

$$Sp = 400 \text{ cm}^2/_2$$

$Tr = 0.98$

$R_n \simeq 0.2$ (Negligible visibility)

I. Compute Ringelmann No (White Plume):

$$R_n = 0.0439 (B_s/B_o) - 8.64 \times 10^{-5} (Bs/Bo)^2$$

$$Bs/Bo = 5.5 \times 10^7 D_s' \omega' / \rho_p'$$

$D_s' = 27$ ft

$\rho_p' = 62$ lb/ft^3

Find ω' (particulate loading, lb/ft^3)

Assume all "excess" vapor condenses suddenly.

$Ts = 50^oC$, $Ta = 80^oF$ (26,7°C)

$$Y'^{o}_{50} = \frac{P^* 50\, m_w}{(P-P^*50)Mg} = 0.0844 \text{ lb water/lb dry gas}$$

$$Y'^{o}_{267} = \frac{P^*26.7\, M_w}{(P-P^*26.7)M_g} = 0.0216 \text{ lb water/lb dry gas}$$

$$M'_c = 0.0844 - 0.0216 = 0.0628 \text{ lb water/lb dry gas}$$

Assume dry gas is air:

Volume = 14,6 ft^3/lb

$$\omega'_{w,o} = \frac{0,0628}{14,6} = 0,00429 \text{ lb/ft}^3$$

$Bs/Bo = 5,5 \times 10^7 \times 27 \times 0.00429/62 > 10^5$

White plume highly visible

$R_n \simeq 4,5$

5. DISCUSSION

The Effects of the Water Spray

The indications are, from the example calculations, that the effects of the water spray on the emissions from test cells are minimal.

The sample calculations indicate a small reduction in plume visibility. The validity of this finding is questionable due to uncertainties of some approximations. One can accept such a finding only if it is assumed that the presence of the liquid improves the coagulation or agglomeration process. Such an assumption is probably valid, but might not be for relatively smaller amounts of excess liquid. This analysis discounts the importance of a "steam" plume.

The sample calculations show that for the example case about one percent of the particulates were retained. The retention of soluble gases was less. There seems to be little question that the relatively insoluble gases (CO, NO and most hydrocarbons) are essentially unaffected by the water spray. The contrary experimental findings of Ref 3 in the case of NO could only be due either to error or to some precondition of the water supply which would make it a better solvent for NO than pure water would be.

Approximately twenty-five percent of the soluble gases (NO_2, SO_3 and soluble hydrocarbons) would be dissolved in the liquid phase for the conditions of the sample calculations. The fate of the liquid phase emissions depends upon atmospheric conditions and is beyond the scope of this study.

Effects of Variables

The variables which have the most effect on test cell operation are:

1. Fuel air ratio, (F/A)
2. Augmentor dilution ratio, (R_d)
3. Water spray rate, (m_v)
4. Maximum allowable temperature, (T)
5. Ambient temperature, (To)

The fuel-air ratio (Mass) is not an independent variable for the cell operation, but is set by the engine operating mode. The ratio varies from about 0.005 at idle to perhaps 0.015 at military, and higher yet in case of afterburner operation. The effect of increasing fuel to air ratio is to increase the exhaust temperature.

The augmentor dilution ratio may or may not be an independent operator variable, although it is certainly at least a design variable within certain limits, One might surmise from Ref 1 that the practical lower limit for the dilution ratio may be about 0,5, Unless some gas cleaning device is to be used, there is no advantage in reducing the dilution ratio, The effect of increasing the dilution ratio is to decrease the augmentor temperature as well as the saturation temperature and to lessen the need for spray water,

In the event free liquid water is present, increasing dilution ratio will decrease the amount of soluble gases absorbed, but will increase the fraction of liquid which will remain in the cell as runoff.

The effect of excess water spray is to cool the exhaust slightly (below the saturation temperature), act as a solvent for some constituents and assist somewhat in the coagulation and agglomeration of the particulate particles. Increasing the amount of excess water will increase the proportion of the soluble gases which are absorbed, but will not materially effect the small fraction of liquid which is retained as runoff from the walls.

An additional consequence of excess liquid water is the emission of a dilute acidic mist which will be of relatively large particles (droplets), up to perhaps 50 microns, This mist will tend to settle close by the test cell, most within 200 to 1000 meters downwind. Whether or not this is undesirable depends upon the proximity of the cell to other facilities.

6. CONCLUSIONS AND RECOMMENDATIONS

It is apparent that the water spray has little effect on the cell emissions aside from causing a fallout of mildly acidic droplets close to the cell. Under many operational and atmospheric conditions, this fallout would be inconsequential, and the only harm would arise if the fallout should cause damage to vegetation or structures in the vicinity.

In the absence of any kind of mist eliminator, no significant reduction of emissions will follow from the use of an excess of water spray over that required to cool the exhaust to approximately 350°F, that temperature being the upper limit to prevent spalling of concrete (Ref 1).

There being no advantage in the presence of the water, the recommended procedure is to use only that amount of water which will cool the exhaust to an acceptable temperature. For many operations, there will be no requirement for cooling water.

The required amount of water to achieve an augmentor temperature of no more than 350°F (~177°C) may be estimated to within five percent by the following:

$$m_w = 20 - 0.064 \left(1 + \frac{1}{(F/A)}\right) (1 + R_d)$$

Where m_w is the mass ratio of water to fuel. The error will be "conservative" (that is, safe from overheating) if the air temperature does not exceed 80°F, and if the heating value of the fuel does not exceed 20,000 Btu/lb (~9.6 x 10^6 j/kg).

According to the above, the example run cited in the sample calculations would have required about 120 gallons per minute of water to cool the exhaust to 350°F, whereas the amount actually used was about 600 gallons per minute.

Recommendations For Future Work

It may be possible to redesign the augmentor so that it would approach the behavior of a venture scrubber while still retaining its required capability of reducing the engine back pressure. If this were possible, and if a demister were added to the discharge stack, the particulate emissions could be reduced dramatically, and the fallout of acidic droplets close to the cell could be eliminated.

It is recommended that studies be made to find out if arranging the augmentor as a venturi scrubber and fitting the discharge with demisters would be a cost effective alternative to devices now being considered for gas cleaning on test cells.

References

1. Robson, F. L., A. S. Kesten and R. D. Lessard, "Analysis of Jet Engine Test Cell Pollution Abatement Methods:, United Aircraft Research Laboratories, AFWL-TR-73-18 (1973).

2. Grems, B. D., and D. F. Naugle, "The Effect of Navy and Air Force Aircraft Emissions on Ambient Air Quality:, USAF Civil Engineering Center, AFCEC-TR-76-56 (1976).

3. "Study of Emissions Discharged From Engine Test Cell", Draft Report, Scott Environmental Technology, Inc., USAF Purch. Req. FY 8952-76-60 102, Mod. P000010 (May, 1977).

4. Naugle, D. F., B. C. Grems III and P. S. Daley, "Air Quality Impact of Aircraft at Ten U S Air Force Bases", 70th Annual Meeting Air Poll. Cont. Assoc., Montreal (1977).

5. "Individual Engine Tests and Model Summary Reports", Scott Environmental Technology, Inc., USAF Contract No F29601-75-C-0046, Phase 2, Teledyne (Dec, 1975).

6. Ibid., Phase 3, Andrews AFB.

7. Ibid., Phase 4, Kelley AFB.

8. Ibid., Phases 6 and 7, Tinker AFB.

9. Ibid. Langley AFB.

10. Lyon, F. T., W. C. Colley, M. M. Kenworthy and D. W. Bahr, "Development of Emissions Measurement Techniques for Afterburning Turbine Engines", General Electric Co., AFAPL-TR-75-52 (1975).

11. Finch, S. P. and A. W. Eyl, "Prediction of Test Cell Visible Emissions". AFCEC-TR-76-47 (1976).

12. Society of Automotive Engineers, Inc., "Aircraft Gas Turbine Engine Exhaust Smoke Measurement." , ARP 1179 (1970).

13. Perry, R. H. and C. H. Chilton, Eds., "Chemical Engineers' Handbook", 5th ed., McGraw-Hill, NY (1973).

14. Fuchs, N. A., "The Mechanics of Aerosols", C. N. Davies, Engl. Trans. Ed., McMillan, NY (1964).

15. Cadle, Richard D., "Particle Size Theory and Industrial Applications", Reinhold, NY (1965).

16. Perry, R. H., C. H. Chilton and S. D. Kirkpatrick, Eds., "Chemical Engineers' Handbook, 4th Ed., McGraw-Hill, NY (1963).

17. Conkle, J. P., W. W. Lackey and R. L. Miller, "Hydrocarbon Constituents of T-56 Combustor Exhaust", SAM-TR-75-8 (1975).

18. Crawford, Martin, "Air Pollution Control Theory", McGraw-Hill, NY, (1976).

19. Halow, I. S. and S. J. Zeek, "Predicting Ringlemann Numbers and Optical Characteristics of Plumes", J. APCA, 23, 8, pp 676-84 (1973).

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

A STUDY OF RELIABILITY OF DEEP WELL ANODES FOR CATHODIC PROTECTION

Prepared by:	John A. Stevens, Ph.D
Academic Rank:	Associate Professor
Department and University:	Department of Chemical Engineering and Material Science Youngstown State University
Assignment:	
(Air Force Base)	Tyndall AFB
(Laboratory)	Civil Engineering Center
(Division)	Engineering Material Directorate
(Branch)	DEMR
USAF Research Colleague:	Capt.
Date:	August 12, 1977
Contract No.	F44620-75-C-0031

A STUDY OF RELIABILITY OF DEEP WELL ANODES FOR CATHODIC PROTECTION

by

Dr. John A. Stevens

ABSTRACT

EDITORS NOTE:

Dr. Stevens became terminally ill during the last week of his ten week study and was unable to formally report his work. However, all data and research results are being maintained at the USAF Civil Engineering Center and will ultimately become part of an Air Force Technical Report to be released by the personnel with whom Dr. Stevens worked. For further inquiry into this research contact the following:

Mr. Thomas F. Lewicki
DEMC/AFCEC
Tyndall Air Force Base, Florida 34203

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

TACTICAL SHELTER HARDENING

Prepared by:	Alvin M. Strauss
Academic Rank:	Associate Professor
Department and University:	Department of Engineering Science University of Cincinnati
Assignment:	
(Air Force Base)	Tyndall AFB
(Office)	Civil and Environmental Engineering Development Office
USAF Reasearch Colleague:	Capt Monte H. Callen, Jr
Date:	August 12, 1977
Contract No:	F44620-75-C-0031

TACTICAL SHELTER HARDENING

by

A. M. Strauss

ABSTRACT

A study was made to determine the maximum extent possible to harden tactical military shelters, in the standard ANSI/ISO configurations 8x8x13 ft and 8x8x20 ft, to provide protection against blast and fragment penetration. A threat survey was performed and a typical threat to the tactical shelter was defined. On the basis of the weight constraints due to mobility considerations it was found that protection against small arms was incompatible with that constraint. Once the threat was defined the study proceeded to investigate the possibility of designing a tactical shelter that would be blast and penetration resistant with no increase in cost or weight. It was determined that this was not practical, because of the myriad of other constraints associated with the design of the next generation of tactical shelters. However, recommendations were made on how to come close to achieving this end. The recommendations include: a) using a single sheet high strength panel design; b) using kevlar sandwich composites; c)using fiber reinforced foams; d) using kevlar reinforced foams; e) using add-on kits; f) redesigning the shelter frame.

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

I do not know that I have anything to say on the subjects more specifically discussed in this report, but I hope I shall not do violence to the spirit of your kind invitation or too much presume on your patience if I shall say a few words on that general subject which you discussed with great clearness....."To be more readily understood I shall use your notation and terminology, and consider the most simple case possible."

Letter from J. Williard Gibbs to Oliver J. Lodge
January 8, 1887

A tactical shelter is a rectangular parallelepiped of dimensions 8x8x20 ft or 8x8x13 ft that is capable of efficient transport by helicopter, rail, ship or truck without incurring any structural degradation. The objective of this study is to take the basic shelter as defined above and design into it a significant amount of blast and penetration resistance. To accomplish this objective a number of different aspects of the shelter and its environments must be studied:

I. Typical Nuclear Threat: We should define the typical blast and projectile threat that a shelter deployed in a tactical situation is most likely to experience. For a nuclear blast, the maximum survivable overpressure (peak pressure) is generally assumed to be 7.25 psi. This is not due to any inherent structural limitations, but is a consequence of the fact that at pressures somewhat higher than 7.25 psi, the shelter will overturn, be pulled loose from its anchoring cables, be blown along the ground by the blast. Thus the higher overpressures present a problem that cannot be solved by a light-weight structure. The reasonability of the 7.25 psi figure as an upper bound is clear. Thus our design objective for blast is established.

II. Typical Non-Nuclear Threat: Let us begin by looking at weapons that can supply overpressures in the neighborhood of 7 psi overpressures. These threats include G.P. bombs, artillery, and rockets. However, even a cursory study of these weapons will show that if they explode close enough to a target (shelter) to deliver peak overpressure of 7 psi, they will easily destroy the shelter by penetration of large amounts of shrapnel. Thus the non-nuclear threat is a ballistic one and to meet it the shelter must be able to resist penetration. For tactical shelter purposes I can define the typical shrapnel threat as a 40 gram fragment impacting the shelter with a velocity of 600 meters/second.

III. Small Arms Threat: The weight of a shelter capable of withstanding small arms fire of, say, 7.62mm ball rounds would be too heavy to effectively be air-transportable. Thus a shelter designed to resist small arms fire would cease to be a tactical shelter, but would be a structure with walls of the same order of magnitude as an armored personnel carrier. Therefore, it is recommended that tactical shelters be moved out of areas where it is subjected to small arms fire, or that revetments be erected around the shelter, or that the shelter be place in a depression.

Now a typical threat has been defined: 7.25 psi Peak Pressure of the Applied Pulse. A 40 gram Fragment with Impact Velocity of 600 Meters/Second.

It is proposed that these two criteria be added to the standard shelter design criteria, and that all future tactical shelters meet these criteria.

Moreover, taking typical threats into consideration it is recommended that all six faces of the shelter meet these criteria. Now is the proper time to bring up precisely what is meant by saying "the tactical shelter meets the 7.5 psi and 40gm at 600 m/sec criteria."

Let us begin with the latter. For a shelter (panel) to meet the 40gm 600 m/sec criterion, a projectile with those characteristics, if it penetrates the panel, must have zero velocity upon entering the shelter.

For a shelter to meet the 7.25 psi criterion it must be capable of meeting all the requirements for certification of a new shelter including:

a. Nine mile per hour railroad humping test.

b. Static loading (75 lb/ft^2).

c. Fording (no water entry).

d. Rain (no water entry).

e. Dust (no dust entry).

f. Thermal insulation including low temperature (-65°F), high temperature (120°F) plus a solar heating (BTU) load.

g. Corrosion resistance (salt fog).

h. Blackout capability.

i. Fungus resistance.

j. RFI, EMI, and EMP protection.

k. Helicopter transport.

l. Stacking test (of shelters), etc.

The reason for this definition is that in experiencing the 7.25 psi blast the shelter will undoubtately exhibit some permanent deformation. This can be tolerated if the shelter is able to perform its function. The reason all the tests should be performed is that there are hundreds of uses for tactical shelters and it is not practical to qualify each one at 7.25 psi. Thus we have assumed that if the shelter meets the design criteria before and after the blast, it can perform its function after the blast.

Among the uses of tactical shelters are: ILS systems, artillery control, antiaircraft systems, radar systems, telephone units, communications systems, helicopter transmission repair facilities, machine shops, photographic development facilities, latrines, CBW shelters, data reduction facility, drone control, operating room, intensive care unit, power station, weather station, and meeting room.

The demand for tactical shelters in the coming years will doubtless increase significantly due to the changing international conditions, which are now beginning to become clear and the type of conflict we are likely to encounter in the future. To begin with, the loss in the number of overseas bases and the degradation (by some foreign control) of the mission of others means that we are no longer as close as we once were to the scenes of prospective military confrontations. Thus, supplies, personnel and material will have to be brought over increasingly larger distances. In addition the required response time to threats has been steadily decreasing. Thus the modern scene is generally one in which there are a few well stocked bases far away from many areas of potential threat and the time required for an effective threat response is in the neighborhood of 24 hours. These problems have been somewhat counter balanced by some efficacious developments in cargo aircraft. They have become faster with greater cargo capacity and the cost of shipping a ton of cargo has concommitantly dropped to a relatively low level. At this point the capability for quick long range response to long distance threats exists. Thus the need for tactical shelter systems is obvious. There is not sufficient time or manpower to dump equipment

at a tactical site and have it set up and put into an operable configuration. The equipment must be preconfigured and in operating order immediately upon delivery. Moreover, the equipment and its operators must be protected from the environmental hazards it will encounter. In addition, the vagaries of warfare dictate that this equipment be mobile to the extent that it can be almost instantaneously moved by truck or helicopter to a new site.

It is clear that shelters of the electronic type, radars, missile control, artillery control, communications, etc. will be high priority targets on an agressor's list. However, given the many uses of the tactical shelter, and the inability, in some cases, to distinguish their functions from their external appearance, all shelters in tactical areas should have the penetration protection specified above.

There are three basic types of shelter construction:

I. Aluminum faced rigid foam with aluminum reinforcing beams.

II. Aluminum faced honeycomb (resin impregnated paper or nomex)

III. Aluminum faced plywood with rigid foam core.

The aluminum faces are typically 0.8mm thick and the foam is approximately 5cm thick. Typically, the foam shelter is reinforced with 3in x 3in aluminum hat sections. The honeycomb shelters have either kraft paper or nomex cores with cell sizes from 0.25 in to 0.40 in with densities of 3.0 to 5.0 lb/ft^2.

It has been known for some time that the standard shelter would not survive the effects of a moderate blast environment. In the 500 ton TNT Dial Pack explosion in July 1970 with S-250, S-335 and S-390 shelters it was found that empty shelters did not survive, but shelters with equipment racks did. The next explosive event (1972) called Mixed Company used S-280 shelters hardened by the addition of aluminum sheet and showed that shelters without hardening or racks failed catastrophically while the aluminum sheet strengthened shelters survived.

The conclusion of the Mixed Company test was that the S-280 shelter must be structurally modified to survive a 7.25 psi blast. In the 1976 Dice Throw event two retrofitted S-280 shelters, a paper honeycomb S-280, a shelter in a revetment, a shelter on a truck in a ditch and an S-280 on a truck (for overturning data) were tested.

The retrofitted (hardened) shelters had bonded to them aluminum honeycomb kevlar faced panels as shown in Figure 3. The kevlar panels were formed of nine layers of fiber bonded with epoxy. The kevlar provided bending stiffness to resist the blast, thermal radiation protection and fragment protection. This retrofit has a weight of 363 kg [1].

Without going into detail it has been determined that the 7.25 psi peak overpressure is survivable by a tactical shelter with a retrofit, and it is possible to build this survivability into the next generation of tactical shelters.

It was determined that expandable shelters will not even survive blast at the 3.0 psi level. Therefore we do not consider expandables in this report, because its construction is inconsistent with blast protection.

The structural analysis of an S-280 shelter subjected to a 5 and 7.25 psi overpressure was performed by the Navy Civil Engineering Laboratory at Port Hueneme CA using the SAP IV code on a quarter panel (used because of symmetry considerations). The S-280 foam and beam standard shelter was analyzed and the standard S-280 shelter was analyzed and the standard S-280 retrofitted with aluminum honeycomb sandwich was similarly analyzed as shown in Figure 4. The bases for the calculations are presented below in the appendix [1].

Another important point relative to Air Force shelters [2] is that tie down cables have a negligible effect on the shelter response except locally at the attachment points. There is in existance a computer program [3] that will predict the overturning of a shelter. According to preliminary calculations; an 8x8x13 ft shelter with tie down cables will overturn at the 10 psi overpressure level.

RECOMMENDATIONS

a. It is possible to design a blast and penetration resistant tactical shelter with small penalties for weight and cost, but with the advantage of having a structure better able to meet many of the other requirements for a tactical shelter.

b. The use of a single sheet high strength panel must be carefully looked into. This panel of, say, boron-graphite fiber reinforced material would have all the necessary strength properties. Equipment could be mounted on the wall itself by means of aircraft adhesives. In fact, if a 32 ft single sheet of the material could be made it could be bent into shape, necessitating the use of three fewer joints. This design has clear advantages, the major one being simplicity: no delamination problems, none of the moisture, fungus or strength problems. However, there will be a thermal problem. The insulation abilities of a single sheet is not as effective as a composite. Therefore an add-on kit of insulation (to be placed on the outside) should be available for use when the situation warrants it.

c. The use of a single sheet of high strength material bonded to a kevlar fiber material should be looked into. Typically, the metal would be the inside layer and the kevlar the outside layer. The kevlar would provide thermal insulation and significant penetration resistance.

d. A single sheet kevlar panel is also a possibility for tactical shelter use. Typically, this panel would have a corrugated or a ribbed construction to increase its stiffness.

e. The use of kevlar in sandwich construction must be looked into with some intensity. The use of kevlar as one or both panels in a honeycomb construction is probably one of the most practical concepts for a tactical shelter panel. We could have two faces of kevlar with an aluminum honeycomb core. One face of aluminum and one face of kevlar with a paper or nomex core. One could also have aluminum sheet glued to kevlar sheet for each face of a sandwich construction. The use of kevlar in conjunction with polycarbonate foam instead of aluminum in the configurations described above should also be investigated, as should the same configurations for kevlar-fiberglas combinations.

f. The use of a foam filled sandwich should not be abandoned. One should look into kevlar reinforced rigid foams and the commercially available glass reinforced polyurethane foam.

In addition there are off-the-shelf syntactic foams with very desirable properties for shelter use. These foams should be configured with many (kevlar, aluminum, polycarbonate, fiber reinforced composites, fiberglas) facing materials for an appropriate evaluation.

g. Another concept that must be carefully studied is that of the add-on kit for blast and penetration resistance. Even though the use of an add-on kit is contrary to the principles of mobility, rapid deployment, and low cost, the concept is useful and the data gathered from it would be valuable in any trade-off analysis.

h. The frame and connections to it should be designed to dissipiate as much energy as possible and maintain the structural integrity of the panels.

REFERENCES

1. CRAWFORD, J.E., et al, Summary of Results for the Stress Analysis of the S-280 Shelter and the Proposed Strengthened S-280 Shelter, TM No. M-51-76-03, Civil Engineering Laboratory, NCBC, Port Hueneme, CA, 1 Jun 1976.

2. CULLIGEROS, J.M., and J.P. WALSH, Finite Element Analysis of the Blast Response of a Complete Shelter System and Effects of Equipment Rack Modifications, BRL Contract Report No. 324, December 1976.

3. HOBBS, N.P., et al, TRUCK-A Digital Computer Program for Calculating the Response of Army Vehicles to Blast Waves, Kaman Avidyne Report No KA TR-136, Burlington, MA, March 1977.

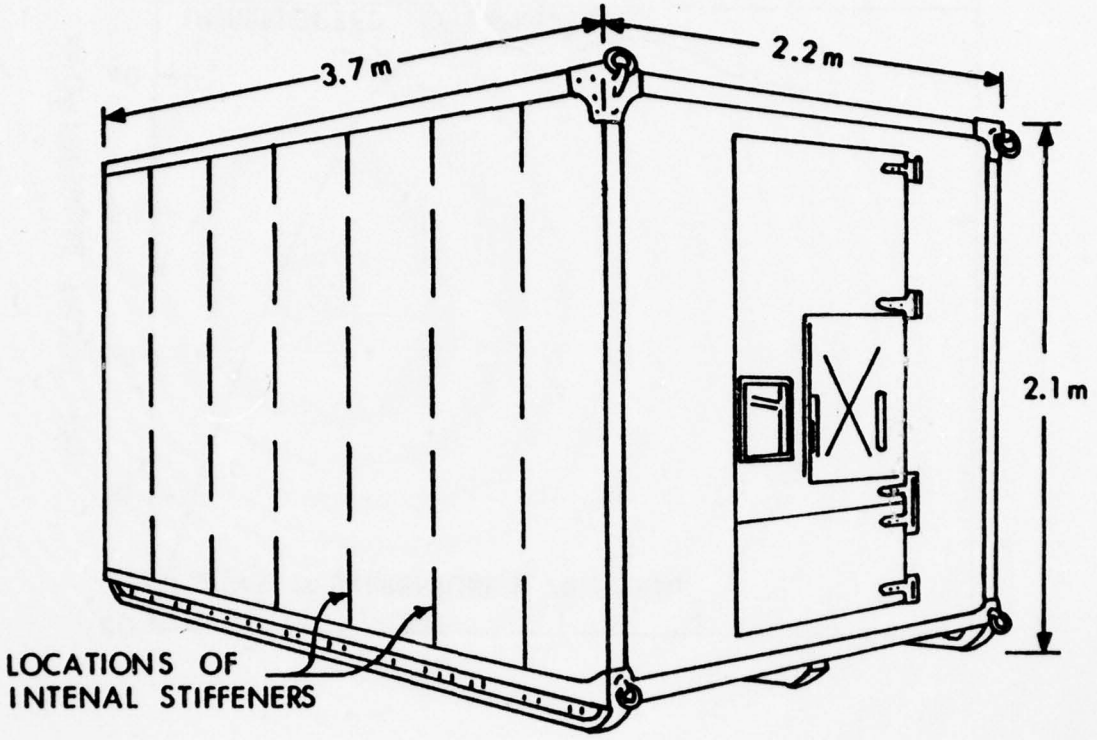


Figure 1. Schematic of S-280 Shelter

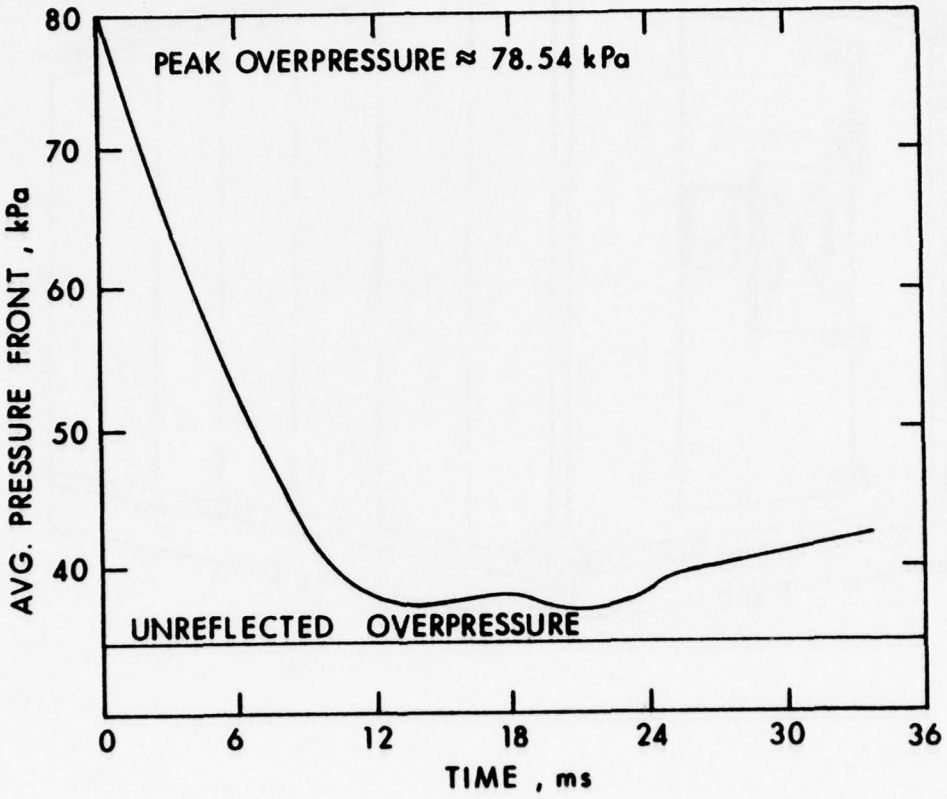


Figure 2. Typical Pressure - Time History

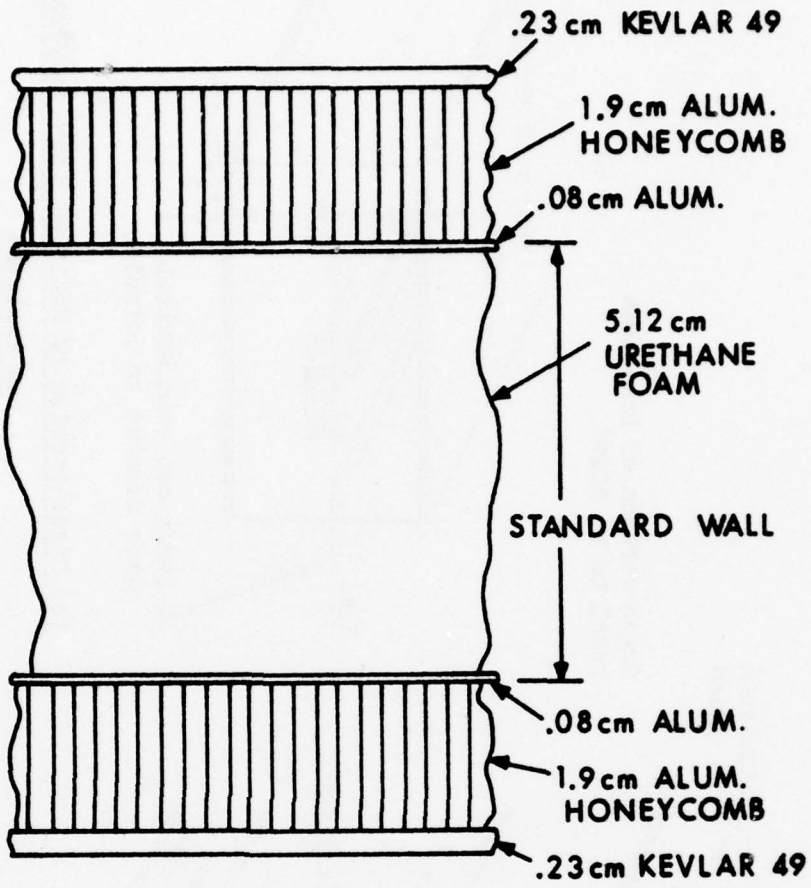
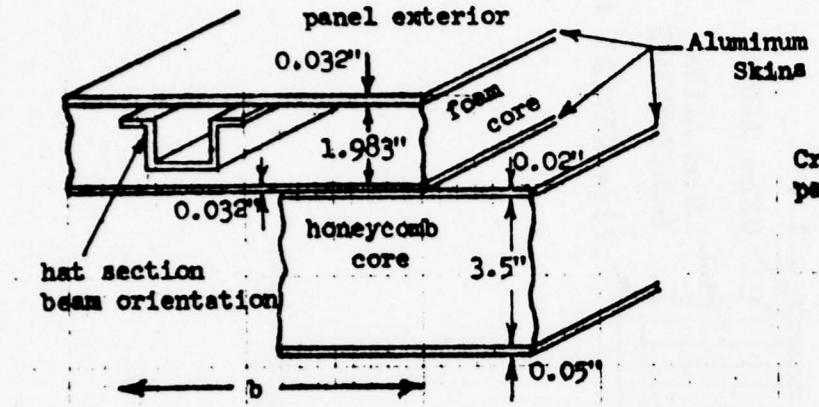
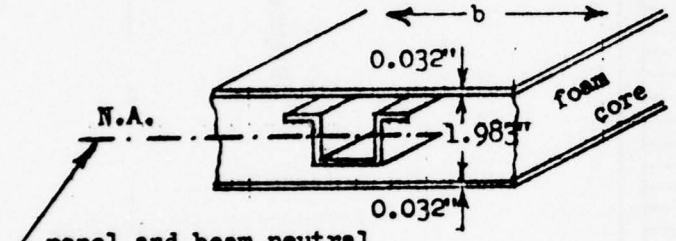


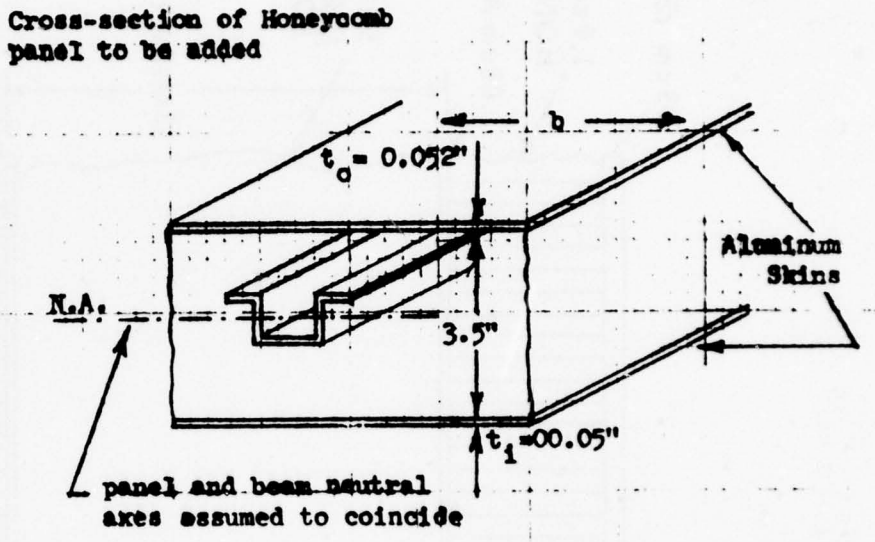
Figure 3. Hardened Wall



(a) Cross-section of Modified S-280 Shelter



(b) Cross-section of S-280 (Al/Foam) Shelter as analyzed



(c) Cross-section of Modified S-280 (Al/Foam/Honeycomb) as analyzed

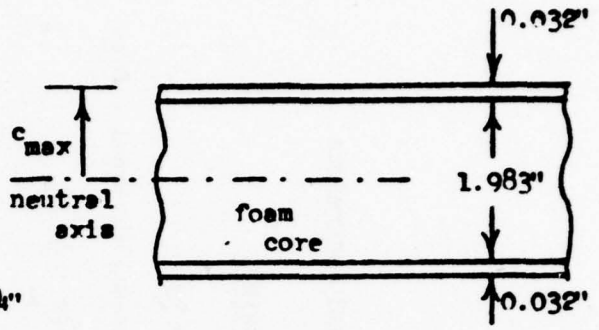
Figure 4 - Cross-sectional Diagrams of S-280 Shelter Model Panels

Bending stress may be computed from the bending moment.

$$\sigma^b = \frac{M\,c}{I}$$

The in plane (membrane) stress may be computed from the in-plane equivalent stress,

$$\sigma^m = \sigma^e \frac{t_e}{t_{Al}}$$



Aluminum/Foam Panel

c_{max} = 0.99 in.

I = 0.065 b $in.^4$

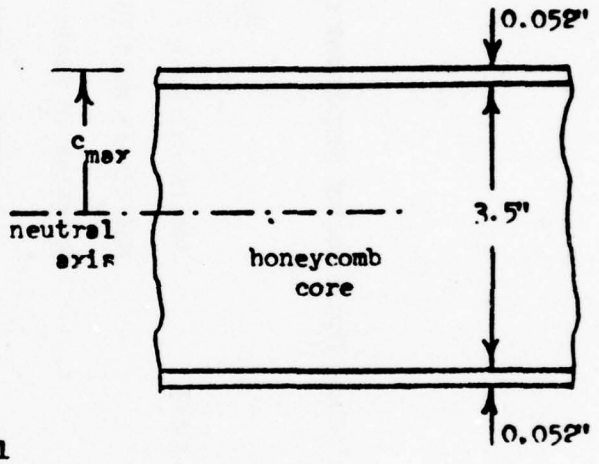
t_e = 3.49" t_{Al} = 0.064"

Therefore,

$$\sigma^b = 15.23\ M \text{ psi}$$

and,

$$\sigma^m = 54.53\ \sigma^e \text{ psi}$$



Aluminum/Foam/Honeycomb Panel

c_{max} = 1.8 in.

I = 0.315 b

t_e = 6.09" t_{Al} = 0.101

Therefore,

$$\sigma^b = 5.71\ M \text{ psi}$$

and,

$$\sigma^m = 60.27\ \sigma^e \text{ psi}$$

Figure 5 - Basis for Panel Stress Calculations

Table I. Dice Throw Design Requirements and Restraints

REQUIREMENTS	RESTRAINTS
Yield - 36 kilo Metric Tons (40 KT)	Additional Mass - 318 kg
Overpressure - 50.4 kPa	Additional Wall Thickness, External - 5 cm
Thermal Radiation - 64 cal/cm^2	Decrease in Aisle Space - 30 cm
Fragments - 0.3 to 3.9 g	

Table II Hardened Wall Material Properties

ALUMINUM FACES	URETHANE FOAM CORE	ALUMINUM HONEYCOMB	KEVLAR PANELS
6061-T6 Alloy	Density = 32 & 64 kg/m^3	Density = 130 kg/m^3	KEVLAR 49 cloth
Modulus of Elasticity = 73 GPa	Shear Modulus = 1.4 & 2.8 MPa	Cell Size = 3.2 mm	Area Density = 0.23 kg/m^2
Yield Stress = 276 MPa		Foil = 0.051 mm	Style = Du Pont 328
		Thickness = 19 mm	Plain Weave
		Type: Hexcel 5056 CR III	Epoxy Resin: Hexcel F-185

BOND MATERIAL	SHOCK ISOLATORS
Hexcel HP-302 - 28.5 MPa lap shear strength	Barry Cupmounts UC-2060-T6
Hexcel HP-326 - 22.4 MPa lap shear strength	
Hexcel HP-347 (KEVLAR to Honeycomb) 40 MPa lap shear strength	

TABLE III.S-280 Shelter Panel Properties

	Al/Foam Panel			Al/Foam/Honeycomb Panel		
	Al Skin	Foam	Equivalent	Al Skin	Honeycomb	Equivalent
t (thickness, in.)	0.032	1.983	3.49	t_o = 0.052 t_i = 0.050	3.5	6.09
E (Young's Modulus, psi.)	1.0×10^7	0	183000	1.0×10^7	0	168000
G (Shear Modulus, psi.)	3.76×10^6	200	69000	3.76×10^6	25600	77700
ν (Poisson's Ratio)	0.33	0	0.33	0.33	0	0.33
w (Weight Density, lb/in^3)	0.0972	0.00116	0.00244	0.0972	0.00301	0.00425
I (Moment of Inertia, in^4)	0.065 b	0	0.065 b	0.315 b	0	0.315 b

Table 4 Results for Dynamic Response of Complete S-280 Shelter

Case No.	Peak pressure of applied pulse (psi)	Peak Displacement δ_{max}(in.)	Beam Element 4 Axial Stress: F / M	σ^F / σ^M	σ (psi)	Plate Stresses due to Membrane Forces and Bending Moments (psi): Element No.	σ_z^e / M_z	σ_z^m / σ_z^b	σ_x^e / M_x	σ_x^m / σ_x^b	σ_{xz}^e / M_{xz}	σ_{xz}^m / σ_{xz}^b	σ_z	σ_x	σ_{xz}	σ
1	5	5.14	1486 21340	2908 90586	93000	3	35 4275	2109 67350	59 7014	3556 110000	2 105	121 1650	69459	13556	1771	11
2	5	1.17	1328 4476	2599 19000	22000	3	25 4330	1507 24724	47 7174	2833 40992	2 141	121 805	26231	43825	926	44
						126	43 1962	2592 11203	24 530	1446 3026	42 1484	2531 8474	13795	4472	11005	21
3	7.25	1.74	2051 6668	4013 28305	32000	3	40 6566	2411 37492	72 10774	4339 61520	3 215	181 1228	39903	65859	1409	66
						126	66 2937	3978 16770	37 819	2230 4676	62 2141	3737 12225	20748	6906	15962	31
4	5	1.55	1380 5082	2700 21573	24000	3	25 1044	1507 5961	46 7381	2772 42146	2 837	121 4779	7468	44918	4900	46
						126	40 1950	2411 11135	18 522	1085 2981	43 1340	2592 7651	13546	4066	10243	20

Case No.

1. Al/foam S-280 Model with imposed rotational symmetry
2. Al/foam/honeycomb S-280 Model with imposed rotational symmetry
3. Case 2 with 7.25 psi pressure pulse
4. Al/foam/honeycomb S-280 Model without rotational symmetry

Table 4. (cont.) Results for Dynamic Response of Complete S-280 Shelter

Case No.	Peak pressure of applied pulse (psi)	Peak Displacement δ_{max} (in.)	Beam Elements 4 Axial Stress: F / M	σ^F / σ^M	σ (psi)	Element No.	σ^e_z / M_z	σ^m_z / σ^b_z	σ^e_x / M_x	σ^m_x / σ^b_x	σ^e_{xz} / M_{xz}	σ^m_{xz} / σ^b_{xz}	σ_z	σ_x	σ_{xz}	σ^p_{max}
5	5	1.33	939	1837	21000	3	21	1266	39	2351	1	60	6468	39026	3920	39000
							911	5202	6423	36675	676	3860				
			4427	18792		126	31	1868	12	723	36	2170	11626	3418	8571	17000
							1709	9758	472	2695	1121	6401				
6	7.25	2.31	2002	3917	36000	3	40	2411	73	4400	3	181	11313	67330	7364	68000
							1559	8902	11021	62930	1258	7183				
			7550	32049		126	63	3797	31	1868	63	3797	21287	6465	15925	31000
							3063	17490	805	4597	2124	12128				
7	7.25	1.98	1377	2694	31000	3	33	1989	63	3797	2	121	9669	58442	5894	59000
							1345	7680	9570	54645	1011	5773				
			6572	27897		126	45	2712	21	1266	53	3194	17113	5349	12267	25000
							2522	14401	715	4083	1589	9073				

ase No.

5. Case 4 with 10% of critical damping
6. Case 4 with 7.25 psi pressure pulse
7. Case 4 with 7.25 psi pressure pulse and 10% of critical damping

δ_{max} - peak deflection (in), node 4, Y-direction

F - Beam axial force (lb)

M - out of plane beam bending moment (in-lb)

$\sigma = \sigma^F + \sigma^M$

σ^e_z, σ^e_x, σ^e_{xz} - computed membrane equivalent stresses

M_z, M_x, M_{xz} - computed bending moments

$\sigma_z = \sigma^m_z + \sigma^b_z$

$\sigma_x = \sigma^m_x + \sigma^b_x$

$\sigma_{xz} = \sigma^m_{xz} + \sigma^b_{xz}$

σ^p_{max} - maximum principal stress

Appendix A CALCULATION OF EQUIVALENT PANEL PROPERTIES

1. For bending and membrane stiffness of Al/Foam

AE: $$\underbrace{bt_eE_e}_{\text{equivalent}} = \underbrace{bt_fE_f}_{\text{foam}} + \underbrace{bt_aE_a}_{\text{aluminum}} \quad (1)$$

EI: $$E_e \frac{bt_e^3}{12} = E_f I_f + E_a I_a \quad (2)$$

since $E_f = 0$, (1) and (2) for a unit thickness (b=1) become

$$t_eE_e = 0.064\ (10^7) = 6.4 \times 10^5 \quad (3)$$

$$E_et_e^3 = 12\,(10^7)\ (0.065) = 7.8 \times 10^6 \quad (4)$$

where $I_a = \frac{b}{12}\left((2.047)^3 - (1.983)^3\right)$

$$I_a = 0.065b$$

then from (3) and (4)

$$t_e = 3.49 \text{ inches}$$

$$E_e = 1.8 \times 10^5 \text{ psi}$$

2. For shear stiffness of Al/Foam

$$t_eG_e = t_fG_f + t_aG_a$$

$$G_e = \frac{1.983\,(200) + 0.064\,(3.76 \times 10^6)}{3.49} = 69{,}000 \text{ psi}$$

3. For mass density of Al/Foam

$$\rho_e = \frac{w_ft_f + w_at_a}{t_e} \cdot \left[\frac{1}{386.4}\right]$$

$$= \frac{0.00116\,(1.983) + 0.064\,(.0972)}{3.49\ \ (386.4)}$$

$$\rho_e = 6.32 \times 10^{-6} \text{ Slugs/in}^3$$

$$w_e = 0.00244 \text{ lbs/in}^3 \text{ (wt density)}$$

4. For bending and membrane stiffness of Al/Foam/Honeycomb

(Note: Figure 4C defines the cross section used for stiffness calculation)

AE: $$\underbrace{bt_eE_e}_{\text{equivalent}} = \underbrace{bt_hE_h}_{\text{honeycomb}} + \underbrace{bt_aE_a}_{\text{aluminum}} \quad (5)$$

IE: $$E_e\frac{bt_e^3}{12} = E_h I_h + E_a I_a \quad (6)$$

since $E_h = 0$, (5) and (6) become

$$t_eE_e = 0.102\ (10^7) = 1.02 \times 10^6 \quad (7)$$

$$E_et_e^3 = 12\ (10^7)\ (0.315) = 3.78 \times 10^7 \quad (8)$$

where $I_a = \frac{b}{12}\left((3.6)^3 - (3.5)^3\right)$

$$I_a = 0.315b$$

then from (7) and (8)

$$t_e = 6.09 \text{ inches}$$

$$E_e = 168{,}000 \text{ psi}$$

5. For shear stiffness of Al/Foam/Honeycomb

$$t_eG_e = t_hG_h + t_a G_a$$

$$G_e = \frac{3.5\ (25{,}600) + 0.102\ (3{,}760{,}000)}{6.09}$$

$$G_e = 77{,}700 \text{ psi}$$

6. For mass density of Al/Foam/Honeycomb

(Note: Figure 4a defines the cross section used for mass calculation)

$$\rho_e = \frac{w_ht_h + w_at + w_ft_f}{t_e}\left(\frac{1}{386.4}\right) = \frac{0.00301(3.5)+0.0972(0.134)+0.00116(1.98}{6.09\quad(386.4)}$$

$$\rho_e = 1.10 \times 10^{-5} \text{ Slugs/in}^3$$

$$w_e = 0.00425 \text{ lbs/in}^3 \text{ (wt. density)}$$

Appendix B CLASSICAL SOLUTIONS

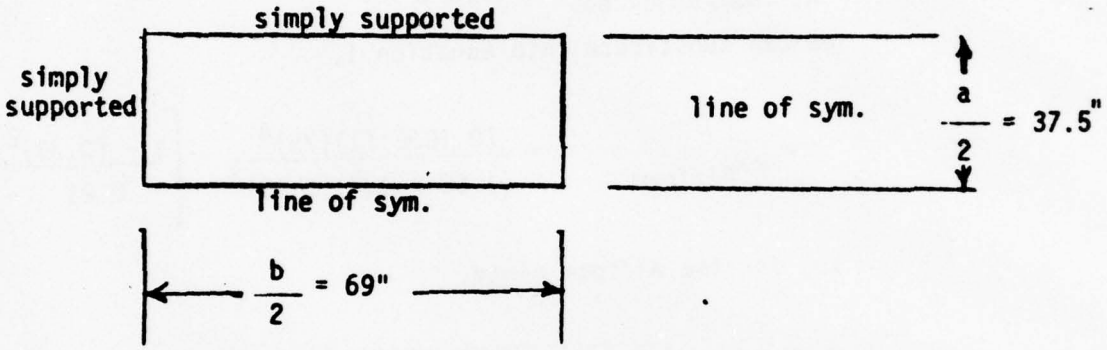
From Timoshenko, Theory of Plates and Shells, p. 129 and Table 5 on p. 133,

for a simply supported rectangular plate $\nu = 0.3$

$$w_{max} = \frac{\alpha\, qa^4}{E\, h^3}$$

our finite element plate analysis uses a $\nu = 0.33$ from pg. 129

$$w_{max} = \frac{\alpha\, qa^4}{E\, h^3} \left(\frac{1-\nu^2}{0.91}\right) \qquad (1)$$



for the Al/foam and the Al/foam/honeycomb plate the aspect ratio is given by

$$\frac{b}{a} = \frac{69}{37.5} = 1.84$$

from Table 5

$$\alpha\Big|_{\frac{b}{a}=1.8} = 0.1017$$

$$\alpha\Big|_{\frac{b}{a}=1.9} = 0.1064$$

the curve for α is approximately linear in the range $1.8 < \frac{b}{a} < 1.9$

So to find a value of α at $\frac{b}{a} = 1.84$ we will linearly interpolate between the given points

Therefore,

$$\alpha = 0.1036$$

For, $q = 1$ psi

$\nu = 0.33$

$a = 75$ in.

$E_{Al/foam} = 1.8337 \times 10^5$ psi

$E_{Al/foam/honeycomb} = 1.6755 \times 10^5$ psi

$h_{Al/foam} = 3.49$ in.

$h_{Al/foam/honeycomb} = 6.0876$ in.

we can substitute into equation 1,

$$w_{max_{Al/foam}} = \frac{(0.1036)(1)(75)^4}{(1.83 \times 10^5)(3.49)^3} \cdot \left[\frac{1 - (0.33)^2}{0.91}\right]$$

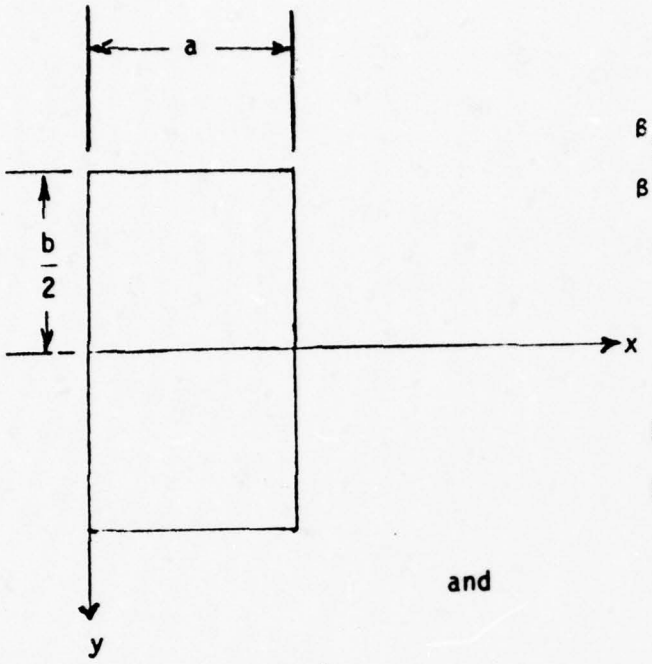
so, for the Al/foam plate,

$$\boxed{w_{max} = 0.412 \text{ in.}}$$

Likewise for the Al/foam/honeycomb plate,

$$\boxed{w_{max} = 0.085 \text{ in.}}$$

Also from Timoshenko, p. 133, we can calculate the maximum bending moments used to compute the in-plane stresses



$\beta_1 = 0.0476$

$\beta = 0.0963$

$M_{x_{max}} = \beta q a^2$

$M_{y_{max}} = \beta_1 q a^2$

and $q = 1$ psi

$a = 70"$

For both the Al/foam and Al/foam.honeycomb plate

$M_{x_{max}} = (0.0963)\ (1)\ (70)^2$

$M_{x_{max}} = 472$ in-lb

$M_{y_{max}} = (0.0476)\ (1)\ (70)^2$

$M_{y_{max}} = 233$ in-lb

1977 SUMMER FACULTY RESEARCH PROGRAM

sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

THE STATUS OF BISHOP PINE (PINUS MURICATA) ON VANDENBERG AIR FORCE BASE AND RECOMMENDATIONS FOR ITS MANAGEMENT

Prepared by:	Paul H. Zedler
Academic Rank:	Associate Professor
Department and University:	Department of Biology San Diego State University
Assignment:	
Air Force Base	Tyndall AFB
Laboratory	AFCEC
USAF Research Colleague:	Maj. R. C. Wooten
Date:	5 August, 1977

The Status of Bishop Pine (*Pinus muricata*) on Vandenberg Air Force Base and Recommendations for Its Management

by

Paul H. Zedler

ABSTRACT

Vandenberg Air Force Base is an area of unique biological importance. Bishop pine (*Pinus muricata*) is one of the interesting species found in abundance on the base. No other publicly owned stands of this species occur south of Monterey. The species represents a biological resource of esthetic, scientific, and possible economic value.

Field studies were conducted to assess the current status of bishop pine on the base, with the aim of recommending what, if anything, should be done to protect the species. All of the areas in which pine occurred were visited, and eight stands were selected for intensive sampling. The sampling included the usual measures of tree diameter, height, and density. In addition, cones were counted on each tree and cone collections made from several trees in each stand. In the laboratory the seeds were extracted from sub-samples of cones and counted, and sub-samples of the seeds taken for estimation of seed soundness. Increment cores were taken to estimate stand and tree age. In addition, data were collected in a pine stand recovering from a controlled burn carried out in 1974.

The results showed that most stands were relatively young and for the most part in good condition. Mortality of trees was predominately that of suppressed trees in dense stands. Most of the stands have seed crops which are large enough to insure that the trees could reestablish in the event of fire. The overall conclusion was that essentially no management was needed in the immediate future, but that the stands should be monitored from time to time to make certain that there is no change in the situation. A series of specific management recommendations are made.

ACKNOWLEDGEMENTS

The study was made possible by a USAF/ASEE Summer Faculty Research Fellowship. Support for field assistants and for some of the field sampling was provided by the National Science Foundation of grant number DEB76-19742 to the author. Many persons provided assistance. Maj. R. C. Wooten USAF/DEV suggested the project, served as technical advisor, and helped in many ways. Mr. Jim Johnston, Base Agronomist at Vandenberg AFB, gave essential logistical support, expedited the necessary paperwork, and shared his knowledge of the base. Mr. J. Fred O'Brien of Auburn University coordinated the fellowship. Bruce McIntyre and Kathleen Guehlstorff, students at San Diego State University, labored under often difficult field conditions to collect data and made many other valuable contributions to the project. Thomas Oberbauer, also a student at SDSU, assisted in the sampling of pine stands on Santa Cruz Island.

During the study I consulted with a number of persons who shared their knowledge of bishop pine or chaparral management with me. Among these persons were: William Critchfield and Gene Conrad, U. S. Forest Service; Richard Minnich and Jonathan Sauer, Dept. of Geography, UCLA; Ken Cole, University of Arizona; Lyndal Laughren, UCSB Santa Cruz Island Field Station; Chuck Rickard, Daly City; Dick McCleery, U. S. SCS; Santa Maria, and J. Everard and D. Fourt, British Forestry Commission.

OBJECTIVES

This study was undertaken to assist the Air Force in the management of the valuable natural ecosystems of Vandenberg Air Force Base. It concentrates on the bishop pine forest, one of the most interesting and significant ecosystem types on the base. The goal of the study was to document the biological significance of bishop pine, especially with regard to its occurrence on VAFB, to determine its present status on the base, to suggest what management goals might be set, and to provide at least preliminary suggestions as to how these goals might be achieved.

The length restrictions for the final report required that the background material reviewing literature and discussing aspects of the biology of bishop pine be omitted. This section will be added as an appendix to those copies of the report sent to interested parties.

INTRODUCTION

Vandenberg Air Force Base (VAFB) occupies one of the most interesting and significant biological regions in the United States. The coastal ecosystems of California preserved on the base are rich in species, diverse in geology, and are superb areas in which to study and come to understand the interactions between organisms and the environment that shape the natural world.

The significance of the base to ecological science derives from several major factors. First, and most important, Vandenberg AFB is one of the few relatively unmolested portions of coastline left in southern California. The bishop pine forest is one of the distinctly coastal ecosystems present on the base. Others are the sand dunes and coastal mesas. Though growth in California is slowing, there is still good reason to suppose that the relative value of the base as an ecological reserve will increase, and not decrease in the future.

But the area presently occupied by Vandenberg was considered biologically significant even before the population explosion along coastal California of the past decades had begun. Many species were first collected and described from locations now on the base, and many rare species have always been restricted to the area of the base and adjacent localities. The importance of the base as a rare plant locality has increased as these adjacent localities have been destroyed or disrupted.

It was also early recognized that the region around Point Conception was a zone of transition from the semi-desert conditions of southern California and the moister, cooler forested regions to the north. This situation is reflected in the presence, on or near the base, of species whose primary range lies much to the north or south. Such unusual occurrence of organisms out of their normal range are important to biogeographers and ecologists studying the climatic control of ecosystems and organisms, and to paleoecologists in interpreting the past history of the earth's biota.

Besides being in a transition zone, Vandenberg AFB also is in an area with an interesting relation to the Channel Islands. Many of the species found at Vandenberg AFB also occur on the Channel Islands, or have closely related forms on the islands. This situation, by no means fully understood at this time, greatly adds to the interest of the region.

Vandenberg AFB was, of course, created to carry out a specific function, and its mission did not include serving as a nature reservation for coastal California. Nevertheless, though the Air Force may not have desired to occupy such a biologically significant area, it is a fact that they do. The Air Force is entrusted with the control and management of a national resource of considerable value. It is fortunate that the mission of Vandenberg AFB does not require large-scale disruption of the landscape, and that, therefore, the conservation of the natural values is not incompatible with the functioning of the base. The Air Force has a clear responsibility to maintain the natural values of Vandenberg AFB to the maximum degree possible.

METHODS

The conclusions and recommendations of this study are based on field sampling and observation of bishop pine on VAFB made during May, June, and July of 1977. Additional sampling and observation of the pine on the base was carried out in August and November 1974, and May 1975. The bishop pine stands on Santa Cruz Island and the Purisima Hills adjacent to VAFB were examined in July of 1977 in conjunction with this study. Other stands of bishop pine have been observed by the author in Mendocino County California, and near San Vicente in Baja California Mexico.

The purpose of the field sampling was to assess the current status of bishop pine on VAFB, and to obtain at least partial answers to the following questions:

1) Is there evidence that the area dominated by bishop pine is shrinking?

2) What proportion of the stands seem to be in a senescent condition and in possible need of burning or other manipulation to prevent their destruction?

3) What is the prospect that wildfires could destroy bishop pine stands?

4) What is the impact of squirrels and other seed predators?

5) Is active management necessary to maintain the bishop pine, and is so, what form should this management take?

To answer these questions, a sampling program which emphasized the intensive collection of data on representative stands was undertaken. As a first step, a map showing all of the known stands of bishop pine on the base was prepared from the vegetation map prepared by T. Oberbauer during a 1974 resource inventory of the base (Coulombe and Cooper 1976). This map proved to be quite accurate, and field checks produced only a few minor corrections. Field reconnaissance of the stands of bishop pine was carried out with the aid of this map. In addition, tree cores were collected, and the density and age of the different areas noted.

Eight stands were selected on the basis of this initial reconnaissance to cover the range of age and density conditions. It was the intent, only partially realized, to pair sparse and dense stands of each age. Sparse stands were defined to be those with less than 40% canopy coverage of pine. The intent of this was to be able to compare the condition and seed crops of these two stand types to evaluate whether a fire would be likely to result in further reductions in density in the sparse stands.

In each stand, canopy cover was estimated, and the slope and exposure determined. The soil type, in general terms, was noted. Stand age was determined from a minimum of three complete and readable tree cores. The location of each stand was marked on a VAFB map, by checking the location against aerial photographs.

Transects were then laid out within the stand, and the slope along the transect recorded. The transects were located away from the edge of stands. This means that samples were not taken of transition areas between stands. While the final selection of the starting point involved pacing a random distance, the overall location of the transects within the stands was not random. Once a starting point was selected, the transect line followed a predetermined compass heading. The beginning and end of each transect was marked with a steel stake of 3/8 inch concrete reinforcing rod.

The transects were usually 30 meters long unless the stand was too small to accommodate this length. In dense stands quadrats were one by two meters in size, with the short dimension along the transect line, and the plot extending one meter to either side of the transect line. In sparse stands the quadrats were 5 meters along the transect line, and extended 15 meters out on either side of the transect line, for a total area of 150 square meters.

In each quadrat the following data were recorded: For each tree in the quadrat the height, (measured with an extendable height pole), the diameters at 10 cm and at breast height (4.5 ft), the condition, and the numbers of cones. The condition data included noting whether or not the tree was alive, the presence of galls or cankers, evidence of bark herbivory, whether or not the tree had a broken top, or any other unusual aspect of the tree that seemed worth recording. The height was estimated as the distance from the ground to the furthest branch tip. For leaning trees, this was measured along the stem rather than vertically.

Stems which were separated below 10 cm height were treated as individuals, though the data also kept track of which individuals were joined together. Stems separated below 4.5 feet were measured as separate individuals for breast height unless such separation was of minor branches. This occasionally required arbitrary decisions. Fortunately, extreme branching of trees was pronounced only in the less dense stands.

Cone counts were made from the ground by use of a three place hand counter. Cones were classified as either 1) immature, 2) mature and closed, or 3) open, severely damaged, or otherwise unlikely to contain viable seed. The intent of the classification was to permit an estimate of the current seed crop carried on the trees. Cone counts were taken for each 10 cm stem, on each tree, including dead trees.

I cannot claim that the cone counts are without error. The main problem was that on larger trees, it was not always possible to see all of the cones. With experience, it is possible to estimate the number of cones in a cluster from a view which shows only a few of them. However, in the main, the cone counts obtained represent minimal estimates. Check counts on trees showed that successive cone estimates could on occasion differ by as much as 25%, though agreement between independent observers was usually much closer, generally within 10%. Not surprisingly, the error tended to be greatest on trees with the most cones. Despite the problems, I feel that the estimates are valid. However, caution must be exercised in making comparisons between stands where the differences are not large.

The shrubs growing with the bishop pine were also sampled, usually in

one meter square quadrats on only one side of the transect line. The diameter at 10 cm of all stems of all species were measured, with data on living and dead recorded separately. Any species which did not regularly reach tree size at maturity (over two meters tall or 10 cm diameter at breast height) was considered to be a shrub. By this definition, coast live oak (Quercus agrifolia) was the only other tree species encountered in the sampling.

Cones were collected in each stand to be used for estimates of seeds per cone and seed viability. In young stands, it was possible to obtain all of the cones from a single individual. In older stands with taller trees, only individual branches could be harvested. In both instances, the cone was taken along with the attached branch. The age of the branch provided an estimate of the age of the seeds in the cone. If possible, at least 25 cones were obtained from each tree, and at least three trees were sampled, providing a minimum sample of 75 cones per stand.

In the laboratory, the branch at the base of the cones was aged, the length of the cones measured, and the condition of the cones noted. Sub-samples of cones were selected for seed extraction. Cones were placed in a laboratory oven at 80°C for two hours to cause them to open. Seeds were then extracted by pounding and manual removal. The total number of seeds was counted and placed in an envelope. A sub-sample of these envelopes was taken for estimates of seed soundness and viability. Seed soundness was defined as the seed being full and with a clear white appearance. The sound seeds were placed in a tetrazolium solution to test viability. This testing did not produce reliable results, and those data are not presented in this paper.

DESCRIPTIONS OF THE STANDS

A brief description of the stands is given in the last part of the appendix. This appendix will be omitted from the official version of the final report, and included only in those copies distributed by the author.

The locations of the stands is shown in figures 1 and 2. The total area sampled, and the number of trees measured are shown in Table 1.

Table 1. Area sampled, and number of trees sampled for the eight stands.

Stand No.	Total Area Sampled (m^2)	No. Bishop Pine Trees Sampled	
		Living	Dead
1	120	134	52
2	1500	169	48
3	1800	88	0
4	60	63	53
5	60	145	113
6	180	72	70
7	180	98	107
8	60	143	37
Total	3960	912	480

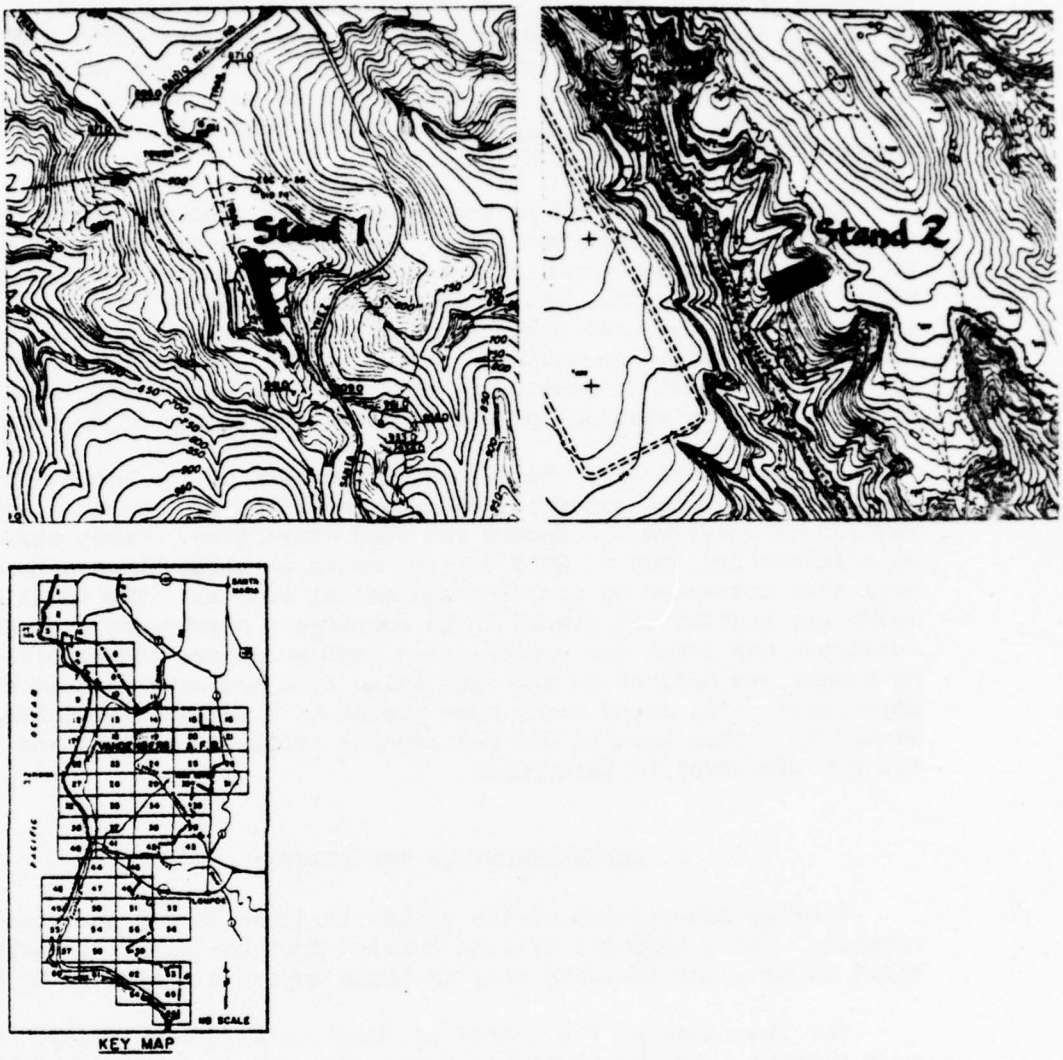


Figure 1. Location of the stands 1 and 2, and a key map for the sheets of the VAFB Master Plan. All of the stands are mapped on portions of sheets of the VAFB Master Plan. Stand 1 is on a portion of sheet 51, and stand 2 on a portion of sheet 38.

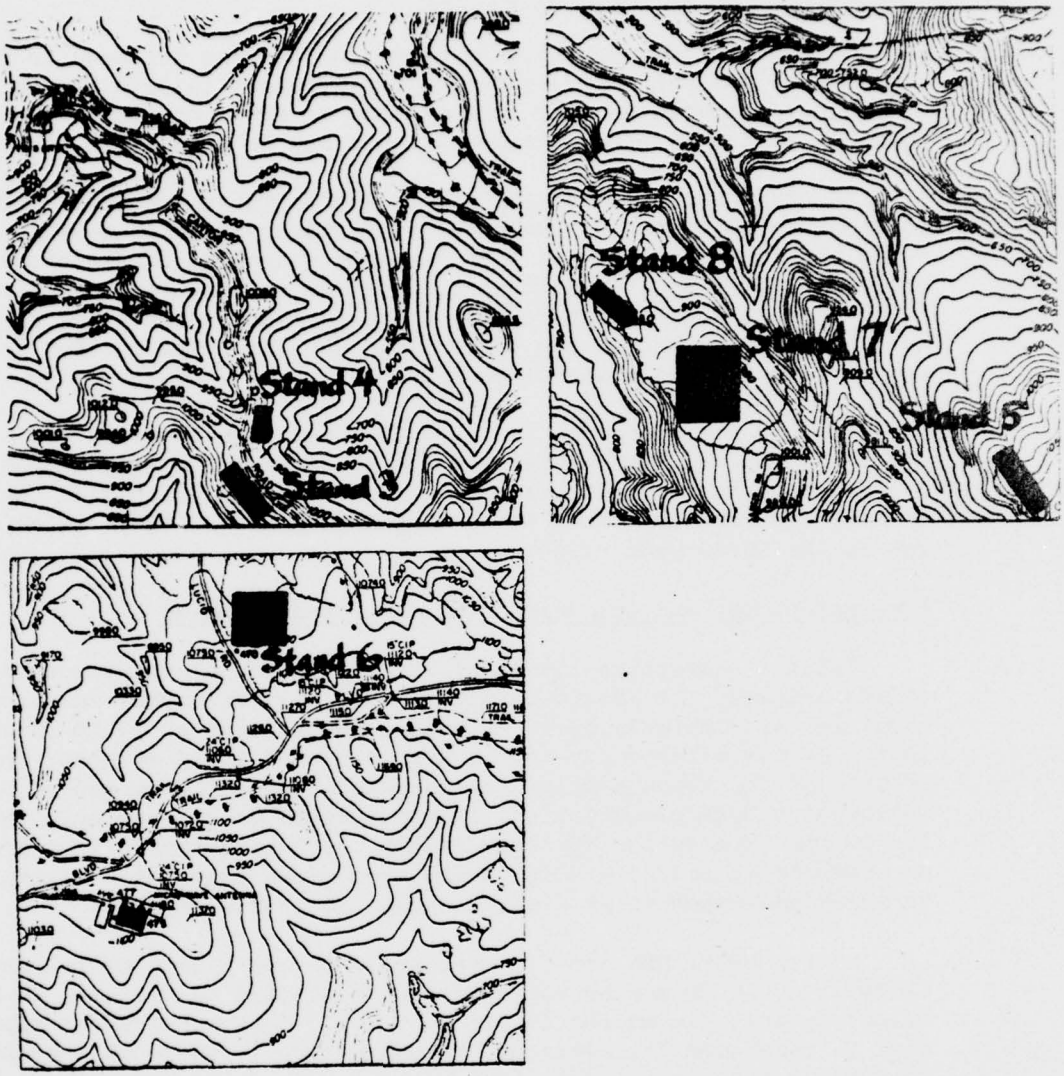


Figure 2. Location of stands 3-8. All stands are mapped on portions of sheets from the VAFB Master Plan. Stands 3 and 4 are on sheet 50, stands 5, 7, and 8 are on sheet 51, and stand 6 is on sheet 55. See figure 1 for a key to the location of sheets.

RESULTS

Distribution of bishop pine on VAFB

The present distribution of pines on the base is shown in Fig. 3. This map is based on the vegetation map prepared in 1974 by Tom Oberbauer, with minor corrections made on the basis of field checking. It is possible that scattered individual trees may be present in other areas.

According to Coulombe and Cooper (1976) there are about 580 acres of bishop pine forest on the base. The stands are best developed on the areas of sandy soil derived from a poorly consolidated sandstone. Many areas with this soil support no bishop pine or only scattered trees. From this indirect evidence it is reasonable to assume that pines might have been more widespread in the past. This idea could be checked for the recent past by examining old aerial photos.

Most of the stands on the base are less than 30 years old, and a number are less than 20. The oldest area on the base is in Pine Canyon. Most of the trees in this area seem to be more than 60 years old. On the whole, the age distribution seems to be tipped to younger stands, with an apparent gap in the 40-60 year range.

Age, structure, and composition of the stands sampled

Table 2 summarizes the major information on stand structure for the stands sampled. It should be pointed out that diameter, and therefore basal area, in this table is based on measurements taken at 10 cm rather than 4.5 feet. As was pointed out in the discussion of field methods, diameters of each of the trees sampled was measured both at 10 cm and at breast height. To simplify data presentation, it was decided to present only the results for 10 cm. The reason for this is that it avoids the serious problems encountered in sparse stands having open-grown trees with multiple stems, in which measurements at breast height are problematical.

Fortunately, DBH and diameter at 10 cm are highly correlated, and diameter at 10 cm may be approximately converted to diameter at breast height by multiplying the former by 0.775. This provides an approximation that is very good for averages, though naturally there will be some error in dealing with individual trees.

The stands of bishop pine on VAFB are not impressive in terms of either the diameter or height of the trees. This is partly because the stands are mostly young, but even in the older stands relatively few trees reach large size. In Stand 2, the oldest stand sampled, the average diameter is only 13.9 centimeters, although there are some large trees (e.g. 41.5 and 32.9 cm at 10 cm height) in the stand.

As is often the case in even-aged forests, there is a close relationship between the density of stems and the average diameter of the stand. This is shown in Figure 4, where the stand density has been plotted against stand average diameter. Despite the considerable range in ages and sizes, there is a clear trend. Young stands with few trees will have a large average diameter because each individual tree will grow faster, and also because diameter growth will be favored over height growth. In old stands,

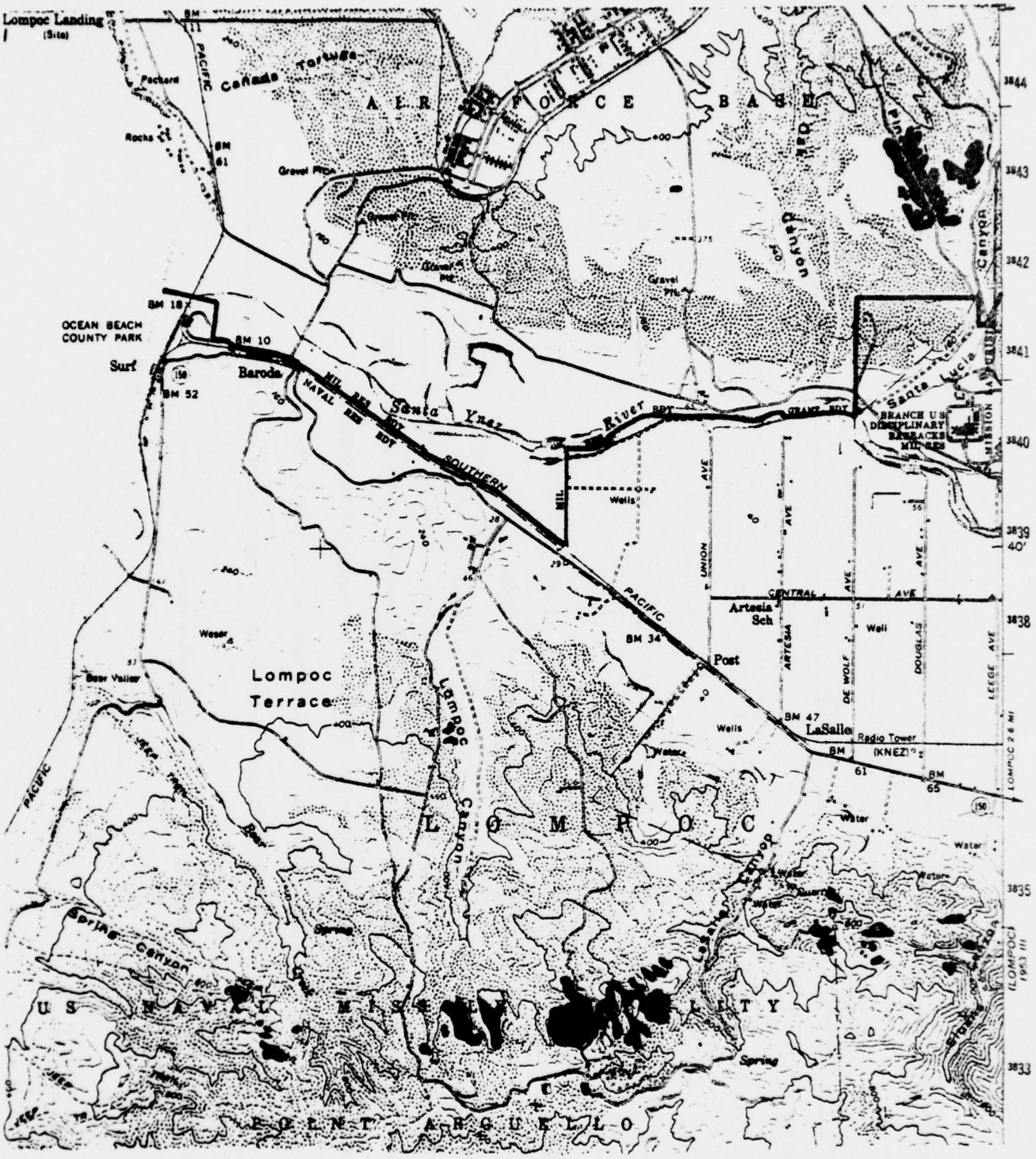


Figure 3. Distribution of bishop pine on Vandenberg AFB. Base map is the USGS 15 minute Point Arguello quadrangle.

Table 2. Major structural features of the stands sampled.

Stand No.	Ages[1]	Density[2] trees/m^2	Ave. Diameter[3] (cm)	Basal Area[4] cm^2/m^2	Ave. Height (m)	Ave. No. Cones[5] per tree	Cones[5]/m^2	Seeds/m^2	Sound Seeds/m^2
1	25	1.12	9.5	52.2	7.22	23.0	11.5	900	754
2	60-72	0.11	13.9	21.4	4.78	5.5	0.6	53	36
3	15 (40-47)	0.05	12.1	7.7	3.16	181.6	8.9	661	507
4	15	1.05	6.6	41.2	3.77	38.3	40.2	3457	3009
5	15	2.42	4.2	43.0	4.85	2.4	5.9	437	354
6	35	0.40	12.9	55.2	7.84	6.7	2.7	73	38
7	20	0.54	10.9	57.6	7.52	15.6	8.5	759	470
8	20	2.38	3.3	29.0	3.11	15.2	36.3	2394	1976

1. Predominant age class. Secondary age classes in parentheses. None of the stands was all-aged.

2. Density counts consider stems joined below 10 cm to be separate individuals.

3. Diameter is measured at 10 cm. An approximate conversion to DBH is to multiply by 0.775.

4. Based on diameter at 10 cm. An approximate conversion to basal area on a DBH basis is obtained by multiplying by 0.6.

5. All cone counts are for apparently healthy, unopened cones.

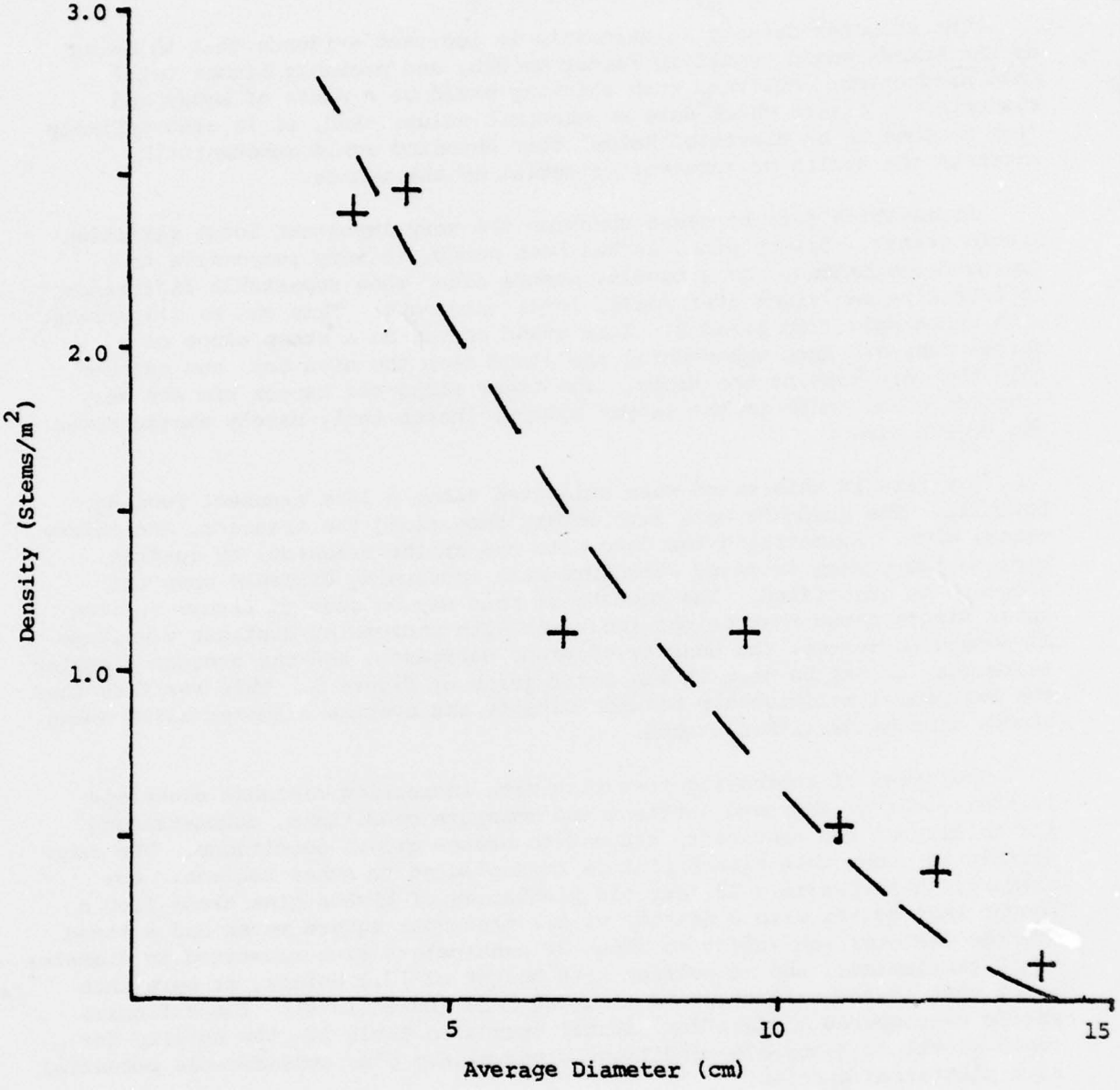


Figure 4. Relationship between stand average diameter and stand density expressed as stems per meter square. Diameter measured at 10 cm above ground level.

the expression of dominance produces fewer but larger trees. In young dense stands, average diameter will be low because of extreme competition. In stand 5, for example, the stems of trees are so close together that it is extremely difficult to walk through the stand. Such stands show that bishop pine is able to survive suppression.

The diameter-density relationship is indirect evidence that thinning of the stands would result in faster growth, and probably higher total wood production. However, such thinning would be a waste of money and resources in stands which have no economic value. And, it is also unlikely (for reasons to be discussed below) that thinning would substantially increase the health or survival potential of the stands.

Summarizing data by stand obscures the very important local variation within stands. Bishop pine, as has been noted, is very responsive to improved conditions. As a result, stands often show remarkable differences in tree size and vigor over small, local gradients. This can be illustrated with an example from stand 2. This stand occurs on a steep slope of a narrow canyon. When approaching the stand from the mesa top, one can see only the very tops of the trees. The trees along the canyon rim are very stunted, while those in the canyon bottom, though tall, barely emerge above the canyon rim.

The data in this stand were collected along a line transect running downhill. The quadrats were five meters deep along the transect, and thirty meters wide. Summarizing the data from one of the transects by quadrat permits the change in stand structure with increasing distance down the slope to be quantified. The results of this may be seen in figure 5. The upper figure shows that height increases with increasing distance downslope. As height increases, the density of trees decreases, and the average diameter increases, as may be seen in the lower graph of figure 5. This verifies that the reciprocal relationship between density and average diameter noted among stands also holds within stands.

The trend of increasing tree size with increasing distance downslope is the result of improved nutrient and moisture conditions, demonstrating the ability of the species to respond to better growth conditions. The same pattern is shown when bishop pine is transplanted to other regions. For example, in Australia, a 23 year old plantation of bishop pine trees from a Lompoc seed source with a density of 0.1 trees per square meter had a stand average diameter equivalent to about 32 centimeters when converted to diameter at 10 centimeters, and an average tree height of 17.1 meters, or more than twice that of the tallest stand recorded here (Doran 1974). These figures should be compared to data for natural stands in Table 2. The ability for rapid growth in favorable conditions gives bishop pine considerable potential as a plantation species.

The condition of the stands may be evaluated by comparing the data on dead trees to the data on living trees. This is done in Table 3. In this table the density, average diameter, and total basal area of the dead trees in each stand is given. The stands in this table are arranged in order of increasing age. Since stands with more live trees will naturally tend to have more dead trees, I have also calculated percentages which express the value for the dead trees in a stand as a percent of the corresponding value for living trees in the same stand. For example, in stand 4 there are 0.88

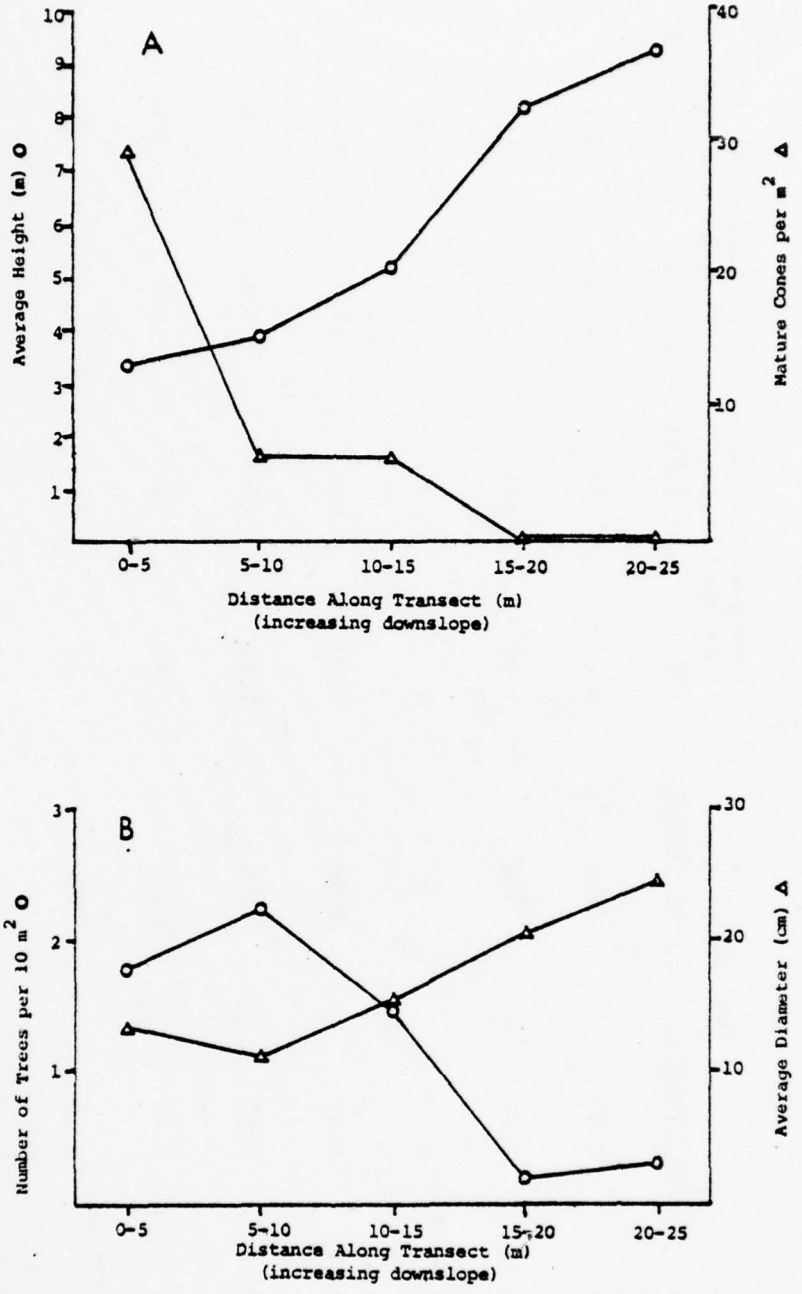


Figure 5. Structural relationships as related to slope position for stand 2 in Pine Canyon. A. Average height and the number of mature cones on an area basis plotted for successive quadrats. B. Number of trees per ten square meters and average diameter plotted for successive quadrats. Diameters were measured at 10 cm above ground.

Table 3. Data for dead trees in the eight stands sampled. Each value is also expressed as a percentage of the corresponding value for live trees in that stand. Example: In stand 8 the density of dead trees is 26% of the density of live trees, that is, there are about four times as many live as dead trees.

Stand No.	Age	Density		Average diameter of dead trees		Basal area of dead trees	
		Dead trees/m^2	% of living	(cm)	% of living	cm^2/m^2	% of living
4	15	0.88	84%	4.24	64%	24.4	59%
5	15	1.88	78%	1.69	40%	5.3	12%
3	15	0	0	0	0	0	0
7	20	0.59	109%	4.20	38%	11.4	20%
8	20	0.62	26%	2.16	65%	2.8	10%
1	25	0.43	39%	3.44	36%	5.2	10%
6	35	0.39	97%	5.27	41%	12.3	22%
2	60-72	0.03	28%	9.16	66%	3.0	14%

dead trees per square meter. This value is 84% of 1.05, which is the value for the density of living trees in stand 4. Large percentage values indicate a situation where there is a high proportion of dead trees.

These data show that there are many dead trees in the stands. However, this is not necessarily a sign of decadence, since it is normal in stand thinning for smaller trees to die as the stand increases in age. This is the case for these stands, since the average diameter of the dead trees is generally well below that of the living trees, indicating that the dead trees tend to be the smaller suppressed individuals. This is borne out by the basal area figures, which indicate that the dead basal area is relatively small in most stands. Stand 3, which is a sparse stand of open-grown trees, had no dead trees at all; which is another indication that the dead trees in other stands are the result of stand thinning.

The exception is stand 4. In this stand dead basal area is 59% of the living, and shows that this stand is suffering excessive mortality. The cause of this apparent decline is unknown, and may be only temporary. For the most part, the stands on the base seem relatively vigorous.

The trees being measured in each of the stands were also checked for visual evidence of galls, beetle damage to stems, or other pests or pathogens. There is a very high incidence of a gall (or canker) which forms large swellings on the branches and results in some malformation of the tree. However, it does not seem very often to cause the death of the tree. Beetle damage was noted, but is rare. Scars where bark has been removed is one of the most common kinds of damage noted. Many of the trees in stands with evidence of squirrel presence (as indicated by nests, characteristic cone refuse, of visual sightings) had trees with numerous large scars on the trunk, generally about 5 by 30 centimeters. These scars most closely resembled porcupine damage, but must be caused by the squirrels eating bark, since there is no other animal other than deer that would be capable of inflicting such large wounds. Since even the largest and healthiest trees in a stand had some of these scars, the scars do not seem to be too damaging.

The relationship between bishop pine and the shrubs with which it grows is an important part of the biology of the species. Table 4 summarizes the shrub data for each stand. The density of shrubs is highly variable among the stands. Not surprisingly, it is highest in the most open stand (stand 3) and lowest in the stand of highest basal area (stand 7). This is consistent with the interpretation that the shrub cover will tend to vary inversely with the competition it receives from the trees. So far as can be determined, bishop pine will always overtop brush, and it is therefore common to find dead shrubs beneath the canopy of bishop pine. Very large open-grown trees with branches to the ground will actually create a conspicuous ring of dead brush beneath their canopy. It is highly unlikely that chaparral shrubs can invade and replace bishop pine provided there is adequate pine seed to repopulate burns. However, shrub species do exert competitive pressure, and if stands were destroyed and there was not immediate reestablishment of the pines, the chaparral would replace bishop pine, and might retard the rate of reinvasion by the pine.

Cone and seed data

A large part of the effort in this study went into counting cones on

Table 4. Shrub density and average diameter in each of the stands sampled. Density is in stems per meter squared. Average diameter is based on measurements at 10 cm and is expressed in centimeters.

SPECIES	STANDS 1 Den.	1 Ave. Dia.	2 Den.	2 Ave. Dia.	3 Den.	3 Ave. Dia.	4 Den.	4 Ave. Dia.	5 Den.	5 Ave. Dia.	6 Den.	6 Ave. Dia.	7 Den.	7 Ave. Dia.	8 Den.	8 Ave. Dia.
Adenostoma fasciculatum					2.88	0.7	0.08	1.0			0.02	1.2			1.30	0.9
Arctostaphylos viridissima	0.94	1.8	0.37	1.0	10.23	1.2	0.67	0.7	0.35	1.3	0.6	1.1			9.17	0.7
Arctostaphylos sp.			1.00	1.0												
Ceanothus impressus									0.03	1.15			0.05	1.9		
C. ramulosus					0.58	1.2									0.17	0.9
Diplacus lompocensis			0.16	0.4												
Heteromeles arbutifolia			0.12	0.5												
Quercus agrifolia			0.12	0.8												
Quercus wislezenii	3.14	1.0					0.03	0.8	1.70	1.1	1.6	1.2	0.07	0.9	0.10	1.5
Rubus spp.													0.03	0.2		
Vaccinium ovatum							1.58	0.7			0.39	0.3				
TOTAL SHRUB DENSITY	4.08		1.77		13.69		2.36		2.08		2.61		0.15		10.74	

trees in the field, and in processing the cones and seeds in the laboratory to obtain cone age, seed number, seed soundness, and seed viability. This emphasis is justified because the potential of the species to regenerate naturally depends upon the presence of a large crop of viable seeds. In order to predict if a stand can regenerate successfully when burned, it is necessary to have some idea of the number of seeds on the tree. Comparison of the number of seeds gives a good idea of the relative health of the various stands.

The basic information needed is the number of seeds per cone, so that the cone counts in the field can be converted to numbers of seeds. Figure 6 is a frequency distribution for number of seeds per cone for all of the cones sampled from all trees in all stands. Considering that this is based on 507 counts, the complexity of the distribution is surprising, and is more evidence for the tremendous variability of cones in bishop pine.

This variation is even more evident when the cone data is broken down by stands, as has been done in figure 7. In this figure one can see very striking differences among stands, suggesting that different stands have significantly different numbers of seed per cone. To test this idea a doubly nested analysis of variance was run on the seed data. The variation in number of seeds per cone of cones within a tree, trees within stands, and among stands was evaluated. The results of this analysis are given in Table 5. These results show that the differences in number of seeds per cone among stands are smaller relative to the differences among trees within stands. In other words, it is the individual trees which show the variation and not the stands. This means that it would be appropriate to use a single soundness value for all stands.

Table 5. Analysis of variance to test for significance of difference in number of seeds per cone among trees within stands and among stands. Two asterisks indicate a difference significant at the 0.01 porbability level. N.S. indicates non-significance. Mean squares and F-values calculated according to Sokal & Rohlf (1968).

Source of variation	Degrees of freedom	Mean square	F
Among stands	7	7986.88	1.028 NS
Among trees within stands	13	6716.73	2.488**
Within trees (error)	132	2699.54	
Total	152	3286.61	

Seed soundness was also highly variable. Surprisingly, it was not possible to show a significant relationship between age of the cone (as determined by ring counts) and soundness of the seeds, a result which contradicts Badran's (1949) work on knobcone and Monterey pine. The oldest cones sampled in this study were 19 years old, so that there seems to be no indication of any significant reduction in soundness for this period of time.

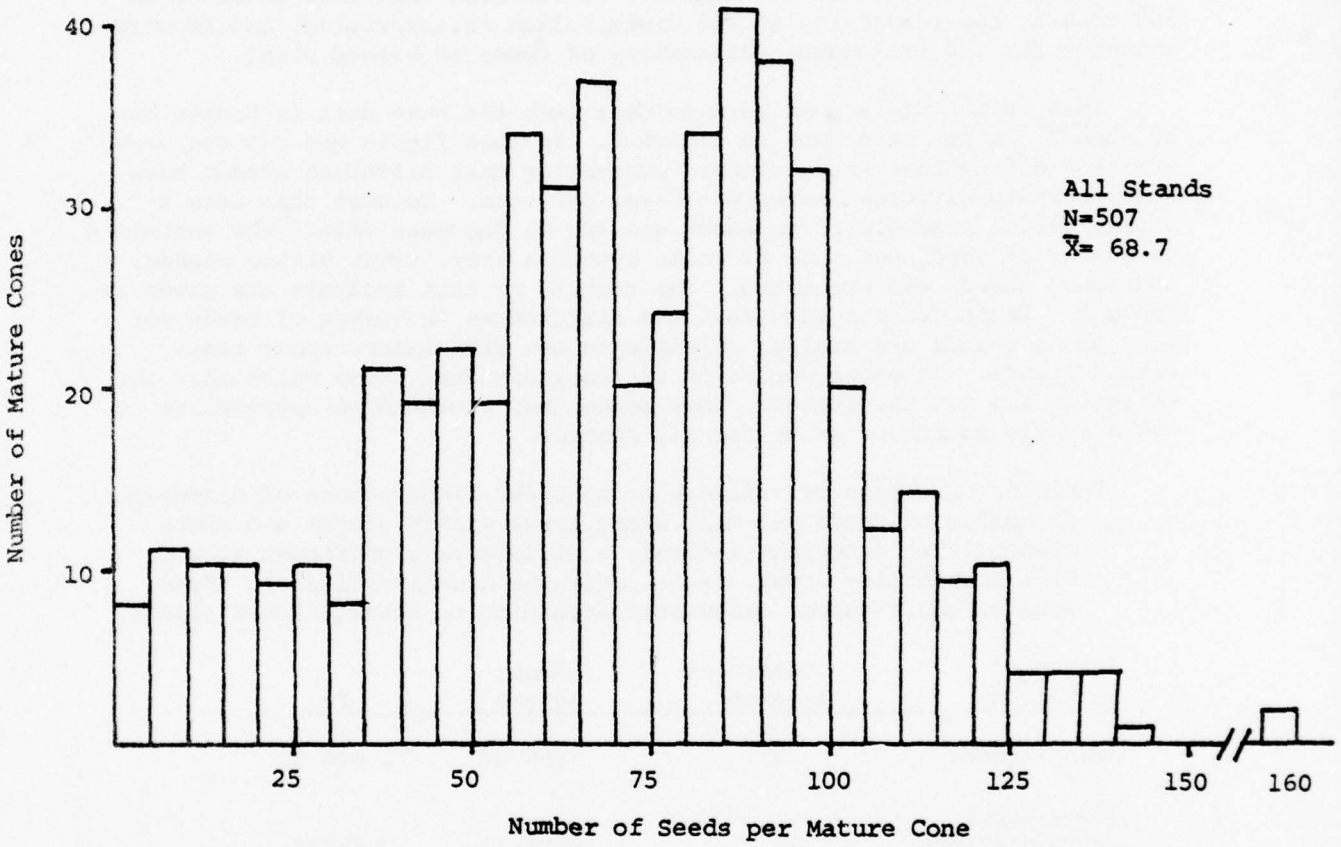


Figure 6. Frequency distribution of number of seeds per mature cone for all stands combined. Number of cones (N) and the average number of seeds ($\bar{X}$) are indicated. Cones plotted here are closed cones which are both mature and free of any type of predation or damage.

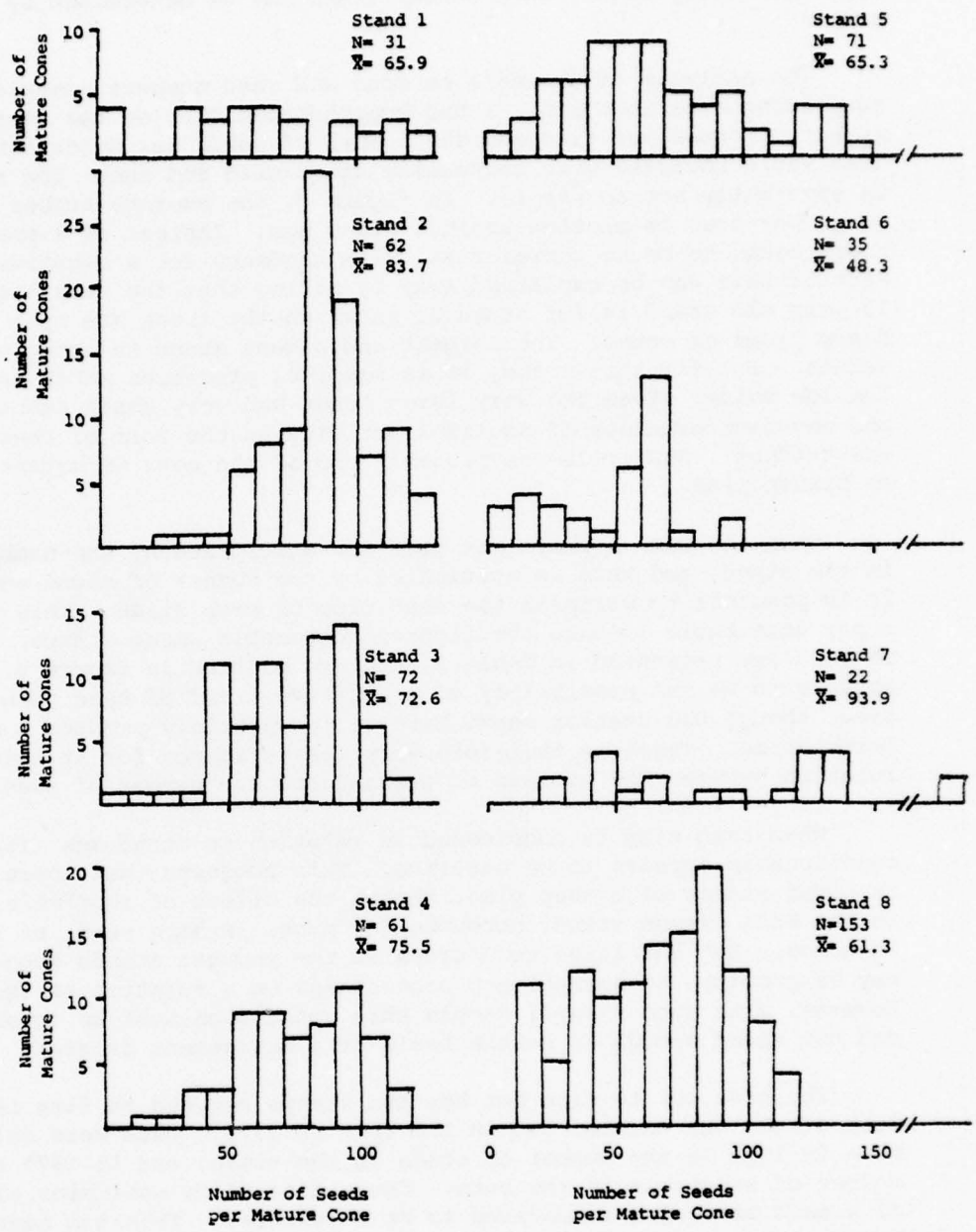


Figure 7. Frequency distribution of the number of seeds per mature cone for each stand. Number of cones (N) and average number of seeds per cone ($\bar{X}$) are indicated. Cones plotted here closed cones which are both mature and free of any type of predation or damage.

Badran (1949) found that there were reductions in soundness with increasing age with knobcone and Monterey pine, and that the estimate of soundness correlated very highly with actual viability as determined by germination tests.

The analysis of the data on cone and seed numbers revealed some surprising relationships. I had hypothesized that on the average the number of cones per tree and the number of seeds per meter square ground area would increase with increasing stand size and age. The relationship is apparently not so simple. In figure 8, the average number of mature cones per tree is plotted against stand age. Instead of a positive correlation, there seems to be no correlation, with tendency for a negative relationship. Part of this can be explained away by noting that the very high value for a 15 year old stand is for stand 3, in which the trees are open grown and bear heavy crops of cones. The largest and oldest stand is among the lowest values. But for this stand, it is squirrel predation which has produced the low value. Even the very large trees had very small numbers of cones, and abundant evidence of squirrel activity in the form of chewed up cones was evident. Squirrels are probably one of the most important influences on bishop pine.

When the cone numbers per tree are multiplied by the number of trees in the stand, and this is multiplied by the number of sound seeds per cone, it is possible to estimate the seed crop of each stand. This was done on a per acre basis to make the figures comparable among stands. The results of this are presented in Table 2, and are plotted in figure 9. There appears to be the possibility of a positive trend of seed numbers with basal area, though the scatter shown here would certainly produce a non-significant correlation. There is, therefore, very weak evidence for at least some relation between the biomass of a stand and the number of seeds.

When seed crop is considered in relation to stand age (figure 9), the relationship appears to be negative. This suggests that there is senescence in older stands of bishop pine, though the effect of squirrels, especially in the Pine Canyon stand, accounts for much, perhaps most, of the reduction in cones. But the large cone crops in the younger stands suggest that it may be possible to burn bishop pine stands on a rotation as short as 25 years. However, with such a small sample this conclusion must be considered tentative, and not sound enough to be the basis of a management program.

The best way to find out how the stands respond to fire is to burn them. Such a burn was carried out in the fall of 1974. Data were collected in the burn in 1974 on the number of cones on the trees, and in 1975 and 1977 on the number of seedlings in the burn. These data allow estimates of the probability of a seed producing a seedling to be calculated. This has been done in Table 6. For two transects the number of sound seeds estimated to be present in 1974 was calculated by multiplying the number of cones per meter square times the average number of sound seeds per cone. Seedlings per meter square was determined from a sample of the same areas taken on 31 May, 1977. Dividing the first into the second gives a number which is an estimate of the probability that a seed will produce a seedling that will survive for two years. These values are surprisingly similar, considering the tremendous number of factors that affect seedling establishment. They are also within the range of values that I have determined for seedling establishment of Tecate cypress in burns in San Diego County. Thus, both the internal agreement and the agreement with independent values suggests that they are reasonable figures.

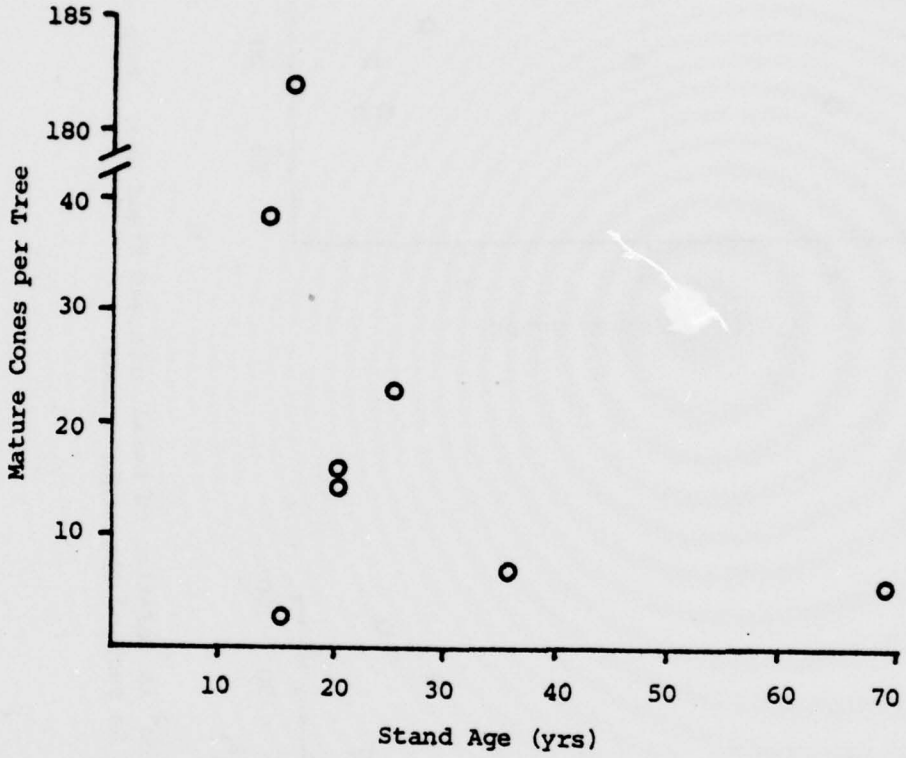


Figure 8. Average number of cones per tree as a function of stand age. Mature cone counts were obtained in the field and include both undamaged mature cones and cones with insect predation.

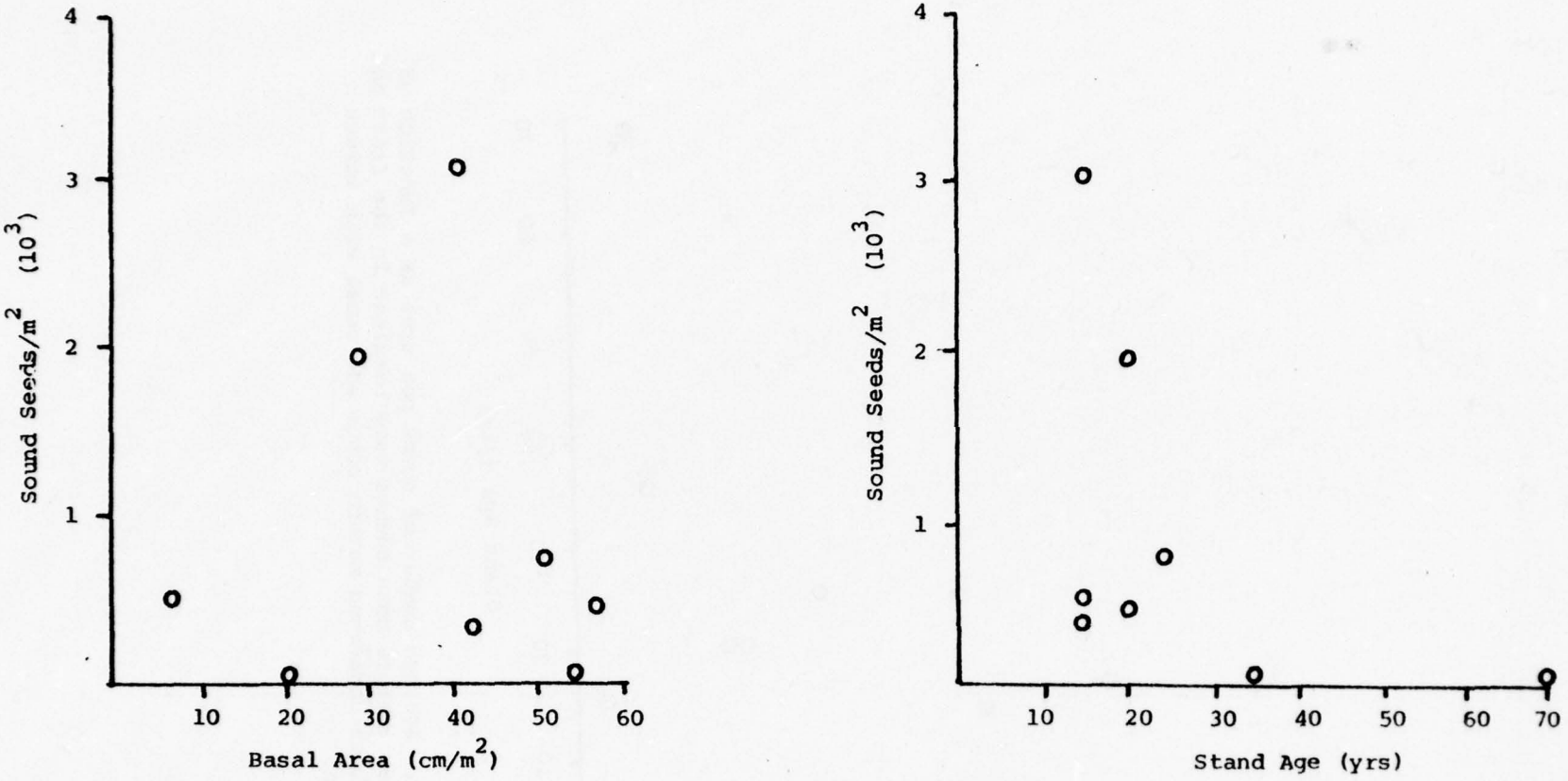


Figure 9. Number of seeds/m^2 in relation to basal area and stand age. Each coordinate represents the average value obtained in each of the eight stands.

Table 6. Seed and seedling data for a controlled burn in bishop pine conducted in fall 1974 and sampled in 1974 and 1977.

Transect	Fall 1974 Estimated sound seeds per m^2	31 May 1977 Seedlings/m^2	Probability of a seed producing a two year old seedling
A	247	0.771	0.0031
B	424	1.925	0.0045

If we take the lower of these as a minimal estimate, it is then possible to predict approximately how many two year old seedlings would be produced in each of the eight stands sampled if they were to be burned with the current crop of cones on them. These predictions are shown in Table 7. This table bears out the conclusion reached on the basis of figure 8. The youngest stands will probably not be the ones most seriously damaged by fire. It is the two oldest stands that are likely to suffer the largest reduction in density, and it may be that a fire would convert them to sparse stands.

Table 7. Predicted densities of seedlings two years after fire for each of the stands, assuming that establishment would be as successful as the low value for the 1974 burn. In this burn 0.31% of the sound seeds produced seedlings still alive two years after initial establishment.

Stand	Estimated two-year seedling density	Condition
1	2.33	dense
2	0.11	sparse
3	1.57	intermediate
4	9.33	very dense
5	1.09	intermediate
6	0.11	sparse
7	1.46	intermediate
8	6.13	very dense

CONCLUSIONS

Significance of bishop pine

It is the conclusion of this study that although bishop pine is not among the species which are rare and endangered, it is, nonetheless, a significant biological resource deserving of protection. This conclusion is based on consultation with other biologists and foresters, as well as

the data gathered during the study. The reasons for this conclusion may be briefly summarized:

1) Although bishop pine is not of commercial significance in the southern part of its natural range in California, including the area on VAFB, the species is grown as a forest plantation tree in Australia, New Zealand, England, and possibly in other countries as well. The natural stands of bishop pine are considered to be valuable sources of genetic material for possible breeding of new strains of plantation trees. Public agencies with stands of bishop pine on their land have an obligation to maintain these for the future.

2) The tremendous genetic variability of bishop pine means that each sub-population of the species is in reality a distinct genetic entity. Therefore,it is not true that preserving the stands of bishop pine in Mendocino County is the same as preserving the entire species. The trees on VAFB are genetically distinct and irreplaceable.

3) Although there is no direct legal recognition of rare and endangered eocsystems in the same way that rare and endangered plant taxa are recognized, all responsible persons interested in the protection of the native biota would agree that species preservation is only accomplished by the protection of ecosystems in which they occur. Furthermore, particular ecosystems are rare enough to deserve special protection, even if they contain species which are reasonably abundant elsewhere. Both of these arguments apply to the bishop pine stands on VAFB. A small number of rare plants are thought to occur in or near the bishop pine stands. In addition, the southern bishop pine forest on coastal terraces is a rare ecosystem type, and deserves protection for this reason.

4) According to Critchfield (personal communication, July 1977), one of the leading authorities on the distribution of forest trees in California, the bishop pine stands on VAFB are the only ones in public ownership south of San Francisco Bay, and therefore are the only large publicly owned areas of the southern or "green" race of the species. Considering the rate of development of this part of the state, it may well be that VAFB is the only place where this race of the species will survive into the next century.

5) Because of the widely accepted view that bishop pine stands are relicts of a once extensive distribution, the stands of bishop pine have unusual biogeographic value. Study of the genetics and ecology of these trees will certainly hold at least part of the answer to the history of vegetation in California.

6) Dense bishop pine forest is the most important habitat on VAFB for the grey squirrel. Protection of the forests will insure that the grey squirrel populations are maintained. The dense and relatively tall forests also provide concealment cover for deer and other wildlife.

7) Last, but by no means of least importance, is the fact that the pine forests are of substantial aesthetic value. Though the trees are not of a kind to delight a production forester, they are picturesque and add a very significant scenic diversity to the base. For at least 26,000 years the trees have grown in the area of VAFB and withstood the stresses of nature, and two waves of human invasion. Our generation should not be the one to destroy these ancient groves.

Present status of bishop pine on VAFB

The data collected during this study indicate that, in general, the stands of bishop pine on VAFB are in reasonably good condition. The majority of the areas are young, and the seed crops carried indicate that there is little danger of extinction should wildfires destroy the stands. Pests and pathogens are present in the stands, and are killing trees, but with the exception of two stands the mortality seems to be primarily falling on the suppressed individuals in dense stands, which is the normal and expected pattern.

Squirrels can have a significant impact on stands. They eat large quantities of seeds, and also evidently eat the bark of young trees. The removal of cones in some instances (e.g. stand 2) is so great that few cones are to be found on the trees except for the new, and probably still inedible, young cones; and old open cones without seed. Duffield (1951) reported that in collecting cones on Camp Cooke (present day south VAFB) he had a difficult time finding enough cones to measure for morphological features. He attributed this to squirrel predation.

The recovery in the 1974 burn is satisfactory, although not nearly so strong as it would have been if there had been a good crown fire. This indicates that control burns, to be most beneficial, should be very intense. An intense fire not only opens all the cones, thus maximizing the release of seed to the soil, but also kills all the trees. In the long run it is probably better to start a bishop pine forest off with healthy seedlings instead of allowing half-dead trees with fire scars to survive.

We could find no evidence that bishop pine stands are actively spreading. If anything the area of bishop pine forest may have receded in the last 50 years, since there seems to be an abundance of habitat suitable for bishop pine on which it does not now grow. It should be made clear, however, that this study produced no direct evidence of stand reduction. The situation should be watched, but it seems that bishop pine is vigorous enough to recover from local reductions in numbers.

It is very evident from field surveys that past activities on the base have not always respected the value of the bishop pine stands. Off the Santa Ynez Ridge Road one of the largest and most interesting stands of young pine was cut into four pieces by a gigantic X bulldozed through it. It may be that there was a compelling reason to do this, but if so, it is not obvious. Other less blatant scars are evident in other stands.

Management recommendations

The recommendations to be made for the management of bishop pine are based on several basic facts:

1) Bishop pine has little, if any, economic potential in native stands on Vandenberg AFB. This does not mean that wood products from the pines could not be sold for cash, but that it would not be possible to manage pine as a commercial species and make a profit.

2) The value of bishop pine as watershed protection, and as natural resource for scientific study and esthetic appreciation far outweighs the meager cash value of the stands.

3) Minimal management is desirable primarily because it will increase the esthetic and scientific value of the pine stands, but it is also the least expensive kind of management.

4) There is at present no crisis with regard to bishop pine. The trend in bishop pine abundance may be downward, but if so, it is at a slow enough rate to provide ample time (decades at least) in which to revise management procedures to forestall extinction.

Given these assumptions I make the following recommendations regarding management of the species.

Recommendation 1: Minimal maintenance. Stands of bishop pine should be unmolested unless there is a very clear and pressing need to disturb them. I do not recommend thinning of stands. Thinning would increase growth, but would destroy much of the scientific interest of natural stands. Similarly, interplanting of trees in sparse stands is neither necessary nor desirable.

Recommendation 2: At present, no special provision for fire protection. I am making the assumption, based on data collected in this study, that if wildfires occurred in the stands sampled recovery would occur vigorously in most cases, and at least with sufficient vigor to prevent extinction in all cases. Traditional methods of fire protection, such as fuel breaks, fire breaks, and drastic thinning are unacceptable because they are expensive and would destroy the values which should be protected. More sophisticated techniques, such as control burning in rotation, are also expensive and, so far as I know, untested for closed-cone species in a coastal environment. Consequently, unless evidence can be found for wildfire drastically reducing the pines, the policy for now should be to take no special precautions.

Recommendation 3: Fire behavior research. The possible value and potential of fuel management for protection of species on VAFB depends a great deal on how fires behave on the base. Since the possibilities for direct experimentation are limited, it may be appropriate to undertake basic research on potential fire behavior. For example, a sampling program could be undertaken to predict fire probabilities and intensities by determining the amounts of fuel in different age chaparral and coastal sage scrub, and field moisture levels at different times of the year. Obviously off-base personnel would have to conduct much of this work.

Recommendation 4: If trees are planted, they should be from seed obtained on-base. A great part of the interest and potential value of bishop pine arises from the tremendous genetic variability it possesses, including the presence of local races. According to William Critchfield, Forest Geneticist with the U. S. Forest Service at the Pacific Southwest Forest and Range Experiment Station, seedlings derived from non-local sources should not be planted among, or even adjacent to, natural stands. This is because of interbreeding which can so confuse the genetics of local populations as to make them useless for scientific study.

Recommendation 5: Establishment of small-scale plantings of bishop pine from seed collected on base. I believe that there is little chance of drastic reduction in pines in the next decade or two. However, it would be wise to maintain a few plantations of bishop pine derived from local VAFB seed as a reserve, in the event that a series of disasters reduces the populations. Since fire is the most likely cause of extinction, the

plantations should be in areas relatively safe from fire. What might be ideal would be to plant several rows or groves of trees in the cantonement area. For added scientific interest, different rows might be selected from different sub-populations on base, or even collected from parents of known characteristics. The plantations would then be scientific experiments as well as a reserve of trees and source of seed. I believe that the U. S. Forest Service research personnel would cooperate in propagating the necessary stock.

Recommendation 6: Briefing of heavy equipment operators. There are many areas of biological interest on VAFB, including the pine stands. A map should be prepared identifying all of the areas with natural values that deserve special protection. All areas of bishop pine should be so classified. Whenever projects or even routine maintenance is planned in areas designated for special protection, the operators of equipment should be briefed regarding proper methods of minimizing impact. Many of the scars on VAFB seem to have been created because sufficient care was not taken.

Recommendation 7: Observation and monitoring. While I have concluded that bishop pine is reasonably safe from drastic reduction, this conclusion must be considered tentative. It will be necessary to monitor the condition of the stands on VAFB from time to time to insure that there is no major change in the situation. Should wildfires occur in the pines, every attempt should be made to obtain information from the burn which could be used to predict the possible fate of other stands. Ideally, the base should immediately let a contract to a qualified ecologist to follow up on the recovery of the stand from fire. Short of this, efforts should be made to alert interested parties to the fire, and to provide persons with the necessary authorizations and logistical support to carry out studies on the burns.

Recommendation 8: A controlled burn might be considered. While at this time there seems to be no need for controlled burning to maintain the stands or for fire protection, it may be advantageous to carry out a test burn, provided that certain problems could be overcome. In particular, I do not believe that fire breaks can be justified unless they encompass a very large area. However, construction or wildfires might isolate a stand which could be burned without the need to engage in major landscape alterations. In such a case, a controlled burn might be considered. Should a controlled burn be undertaken, the provisions of recommendation 7 would apply.

BIBLIOGRAPHY

Axelrod, Daniel I. 1967. Evolution of the Californian closed-cone pine forest. In, Proc. Symposium on the Biology of the California Islands. Ralph N. Philbrick, ed. p. 93-149, Illus. Santa Barbara: Santa Barbara Bot. Gard.

Axelrod, D. I. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. Ann. Missouri Bot. Gard. 62: 280-334.

Badran, O. A. 1949. Maintenance of seed viability in closed cone pines. M. S. thesis. University of California [Forestry], 63 pp.

Cobb, F. W. and W. J. Libby. 1968. Susceptibility of Monterey, Guadalupe Island, Cedros Island and Bishop pines to Scirrhia (Dothistroma pini) the cause of red band needle blight. Phytopathology 58: 88-90.

Coulombe, H. N. and C. F. Cooper. 1976. Ecological assessment of Vandenberg Air Force Base, California. Vol. I Evaluation and recommendations. AFCEC-TR-76-15 Air Force Civil Engineering Center, Tyndall Air Force Base, Florida. 189 pp.

Coulombe, H. N. and C. R. Mahrdt (eds.) 1976. Ecological assessment of Vandenberg Air Force Base, California. Vol. II Biological inventory 1974/1975 AFCEC-TR-76-15 Air Force Civil Engineering Center, Tyndall AFB, Florida. 201 pp.

Critchfield, W. B. and E. L. Little, Jr. 1966. Geographic distribution of the pines of the world. U.S.D.A., Forest Service, Misc. Publ. 991, Washington, D. C. 97 pp.

Critchfield, W. B. 1967. Crossability and relationships of the closed-cone pines. Silvae Genet. 16: 89-97.

Doran, J. C. 1974. Variations in growth of Pinus muricata provenances and comparison with Pinus radiata. Aust. For. Res. 6: 19-24.

Duffield, John W. 1951. Interrelationships of the California closed-cone pines with special reference to Pinus muricata D. Don. Ph.D. thesis on file at U. C. Berkeley. 144 pp.

Epling, Carl and William Robinson. 1940. Pinus muricata and Cupressus forbesii in Baja California. Madrono 5: 248-259.

Everard, J. E. and D. F. Fourt. 1974. Monterey pine and bishop pine as plantation trees in southern Britain. Quarterly Journal of Forestry 68: 111-125.

Fielding, J. M. 1961. Provenances of Monterey and bishop pines. Commonw. For. Bur. Bull. 38. 30 pp. Canberra.

Fielding, J. M. 1961. The pines of Cedros Island, Mexico. Austral. Forestry 25: 62-65.

Forde, M. B. and N. M. Blight. 1964. Geographical variation in the

turpentine of Bishop pine. N. Zealand Jour. Bot. 2: 44-52.

Fowells, H. A. 1965. Silvics of the forest trees of the United States. Agric. Handbook 271, Forest Service, U.S.D.A., Washington, D. C. 762 pp.

Griffin, J. R. and W. B. Critchfield. 1972. The distribution of forest trees in California. Berkeley, Calif. Pacific SW. Forest & Range Exp. Stn. 114 p., illus. (USDA Forest Serv. Res. Paper PSW-82).

Hale, George O. 1941. A survey of the vegetation of Cedros Island, Mexico. M. A. thesis, Univ. Calif. Los Angeles. 96 pp., illus.

Hoover, R. F. 1970. The vascular plants of San Luis Obispo County, California. Univ. of California Press. Berkeley, 350 pp.

Howell, J. T. 1941. The closed-cone pines of insular California. Leaflets of Western Botany 3: 1-8.

James, R. N. and R. L. Knowles. 1973. A timber grade study of 'old crop' Muricata pine. In Report of Forest Research Institute for 1972, New Zealand Forest Service. 38 pp.

Jepson, W. L. 1910. The Silva of California. Memoirs of the University of California Vol. 2, The University Press, Berkeley, CA. 480 pp.

Lindsay, A. D. 1932. Bishop pine (Pinus muricata D. Don) in its native habitat. Aust. Comm. For. Bull. No. 11. 20 pp.

Linhart, Y. B., B. Burr and M. T. Conkle. 1967. The closed-cone pines of the northern Channel Islands. In Proc. Symposium on the Biology of the California Islands. Ralph N. Philbrick, ed. p. 151-177, illus. Santa Barbara: Santa Barbara Bot. Gard.

Little, Elbert L., Jr. and W. B. Critchfield. 1969. Subdivisions of the genus Pinus (pines). U.S.D.A. Forest Service, Misc. Publ. 1144.

Lotan, J. E. 1968. Cone serotiny of lodgepole pine near Island Park, Idaho. U. S. Dept. Agr., Forest Serv., Intermountain Forest and Range Exp. Sta., Ogden, Utah. 6 pp., illus. (U.S. Forest Serv. Res. Pap. INT-52).

Mason, H. L. 1949. Evidence for the genetic submergence of Pinus remorata In Glen L. Jepson, Ernst Mayr, and George Gaylord Simpson eds. Genetics, Paleontology, and Evolution. pp. 356-362. Princeton, N. J.

McMillan, C. 1956. The edaphic restriction of Cupressus and Pinus in the coast ranges of California. Ecol. Mono. 26: 177-212.

Metcalf, W. 1921. Notes on the Bishop pine (Pinus muricata). J. For. 19: 886-902.

Miron, N. T., E. Zavarin, K. Snajberk, and K. Costello. 1966. Further studies of turpentine composition of Pinus muricata in relation to its taxonomy. Phytochemistry 5: 343-355.

Munz, P. A. 1974. A flora of southern California. Univ. of California Press, Berkeley. 1086 pp.

Munz, P. A. and D. D. Keck. 1959. A California Flora. University of California Press, Berkeley. 1681 pp.

Righter, F. I. and J. W. Duffield. 1951. Interspecies hybrids in pines. Jour. Hered. 42: 75-80.

Shelbourne, C. J. A. 1973. A second look at Muricata pine. In Report of Forest Research Institute for 1972, New Zealand Forest Service. pp. 24-25.

Shelbourne, C. J. A. 1974. Recent investigations of wood properties and growth performance in Pinus muricata. New Zealand Journal of Forestry 19(1): 13-45.

Smith, C. F. 1976. A flora of the Santa Barbara region, California. Santa Barbara Museum of Natural History, Santa Barbara, Ca. 331 pp.

Sokal, R. R. and F. J. Rohlf. 1968. Biometry. W. H. Freeman and Co., San Francisco. 776 pp.

U. S. Forest Service. 1974. Seeds of woody plants in the United States. Agriculture Handbook 450. Forest Service, USDA, Washington, D. C. 883 pp.

Vogl, R. J. 1973. Ecology of knobcone pine in the Santa Ana Mountains, California. Ecol. Mono. 43: 125-143.

Westman, W. E. 1975. Edaphic climax pattern of the pygmy forest region of California. Ecol. Mono. 45(2): 109-135.

Whitmore, T. C. 1975. Tropical rain forests of the Far East. Clarendon Press, Oxford. 282 pp.

APPENDIX

THE BIOLOGY AND SIGNIFICANCE OF BISHOP PINE

Taxonomic relationships of bishop pine

Bishop pine (Pinus muricata D. Don) belongs to the subsection Oocarpae of the genus Pinus according to the classification scheme of Little and Critchfield (1969). This group includes Monterey pine (P. radiata) and knobcone pine (P. attenuata) in California and four species of pines found in Mexico. All of these species are characterized, to a greater or lesser degree, by cone serotiny, that is, by having cones which do not open upon maturing, but rather remain closed for many years, opening readily only when heated by fire. For this reason the species in this group are commonly known as "closed-cone pines." This closed-cone trait also occurs in other less closely related species of pines.

The bishop pine is most closely related to Monterey and knobcone pine, but is readily distinguished from them by its two-needle clusters. It is particularly close to the Monterey pine and is reported to hybridize with it (Mason 1949). These two species have almost certainly recently diverged from a common ancestor.

Bishop pine is one of the most variable species of pine anywhere in the world. This variation is present both within and among populations. A large part, probably the greater part, of the variation even within a single stand, is the result of genetic variation. This is most evident in the cones, which vary strikingly from one tree to the next.

Cone differences are also quite pronounced between the island populations on Santa Cruz Island, and those on the adjacent mainland, as on Vandenberg AFB. The island populations tend to have more nearly symmetrical cones and very reduced spines on the cone scales. The cones also tend to be borne at right angles to the branches. The mainland populations, as at Vandenberg AFB, tend mostly to have the cones reflexed (bent down along the branch) with stout spines on the outward part of the cone, producing a very assymetrical cone. However, both of these types occur both on the mainland and on the islands, and intermediates are present at both places.

Mason (1930, 1949) described the island form as a separate species, calling it Pinus remorata. Both Munz (Munz and Keck 1969, Munz 1974) and Smith (1976), standard references for the flora on Vandenberg AFB, retain Pinus remorata. Other authorities (Hoover 1970, Duffield 1951) consider that P. remorata is only a form or variety of P. muricata. This confusion means that Vandenberg AFB has either one or two species of pines, depending on which authority is followed. In this study, care was taken to note the remorata type when it was encountered. Cones like those of remorata were found on trees in the Pine Canyon stand. However, I believe it is most appropriate to consider all of the trees on VAFB to be P. muricata.

The study of morphology (Duffield 1951), turpentine chemistry (Forde and Blight 1964, Mirov et al. 1966) and crossability of the sub-groups of bishop pine (Critchfield 1967) have all led to the recognition of geographical races. According to Mirov et al. (1966) there are at least three chemical

races of bishop pine. A northern race which occurs north from Sonoma County, a central race from Sonoma County to Monterey County, and a southern race from San Luis Obispo County south. The island population on Santa Cruz and Santa Rosa are also said to be distinct, and form a sub-type within the southern race. The northern and southern chemical races correspond to the "blue" and "green" geographical races (respectively) established on morphological grounds.

The northern blue race has bluish foliage, reaches larger sizes and greater ages (as much as 250 years), grows straighter, and tends to have less pronounced closed-cone behavior. The southern trees are shorter lived, smaller, much more branched, and generally have cones which do not open readily except after fire. Very surprisingly, these two races within the same species are extremely difficult to hybridize (Critchfield 1967). Pines are known for the ease with which they form inter-specific hybrids. The failure of crosses within a species is therefore very unexpected. On the basis of this evidence the southern race of bishop pine must be thought of as approaching a separate species. From this point of view, it is one of the rarest trees in the United States.

Paleoecology of bishop pine

Bishop pine or species closely related to it are well represented in the fossil record in California, indicating a long history of these species in the region. Axelrod (1967, 1975) provides the most complete discussion of the prehistory of bishop pine and the closed-cone forest generally. According to him (Axelrod 1967) the lineage of bishop pine extends at least into the middle Pliocene (roughly four million years ago), and very likely back to the early Miocene (roughly twenty million years ago). The immediate ancestor of bishop pine has been given the name Pinus masoni. Bishop pine may have evolved in the uplands of the interior, and then migrated to the coast as the climate became drier. However, the more recent fossils were found at sites and in sediments that indicate that bishop pine has grown along the coast for a long period of time. Fossil cones of bishop pine, including those of the remorata type, have been found along with Monterey pine cones near Point Sal just north of VAFB. These cones, dated at about 26,700 years before the present, prove that bishop pine has had a long history in the vicinity of Vandenberg AFB.

Some of the fossil localities for bishop pine and other closed-cone conifers are at places (e.g. Ventura, Rancho La Brea, San Pedro) quite distant from the nearest present-day stands. This, plus the relative abundance of pine fossils, has led Axelrod (1967) to suggest that the closed-cone pine forest was more widespread in former times. According to him, a "continuous closed-cone pine forest (may have) blanketed the outer coast and islands . . . as recently as 14500 to 12000 years ago" (Axelrod 1967). From this point of view, the present stands of bishop pine are regarded as relicts of a formerly widespread type.

It is also possible that the bishop pine stands might have been considerably more widespread in recent centuries than they are at present. While it is unlikely that the area of Vandenberg AFB was a continuous pine forest two hundred years ago, the impact of settlement may have reduced the extent of the stands. The early settlers in the region might well have been attracted to the bishop pine as a nearby source of softwood lumber or

posts. Heavy grazing can also be very destructive to pine stands, as may be dramatically seen today on Santa Cruz Island. Localized cutting plus frequent fire and grazing might well have greatly diminished or exterminated stands in some areas. On Vandenberg AFB there seem to be many sites which could support bishop pine but which have only a scattering of pines or none at all. It may be that these areas were pine forests in pre-settlement times. Without better historical information or further observation and experiment this must be considered only a possibility.

Ecology of bishop pine

Site preference. It is obvious from even casual inspection that bishop pine does not occur randomly in the landscape. At the broad geographical level, it is always found close to the coast, and apparently depends in some way upon the moderating influence of the ocean. Most likely it requires merely that cool temperatures, overcast skies, and frequent fogs reduce transpiration below that which prevails further inland. There is also a marked tendency for stands to be best developed, or in some cases limited to north-facing slopes, another circumstance suggesting that they are intolerant of extreme dessication. It may also be that the species is frost-sensitive, though it seems likely that most stands must experience at least occasional periods below freezing. However, as with some other plants (e.g. saguaro cactus) it may be that the species is intolerant of prolonged freezing, a condition unlikely in a coastal location. Shelbourne (1974) reports that bishop pines from several localities along the coast were planted south of San Francisco for an experiment, and were subjected to seven to ten days of freezing temperatures with a minimum of -8.3°C. All of the bishop pines, except those from Humboldt and Mendocino County, suffered some damage.

While in any small area the bishop pine seems to show very strong soil preference, it definitely cannot be said that the species is severely restricted with regard to ecological substratum or soil type. On Vandenberg AFB it is most abundant on young, poorly consolidated sandstones, but it also occurs on shales as at Pine Canyon and in the nearby Purisima Hills. On Santa Cruz Island there are populations on three distinctly different geological substrates, including a volcanic formation. Plantations in England, New Zealand, and Australia, and individual specimen trees or small experimental groves planted at various places in California also show that bishop pine can grow on many different soil types, and tolerate different climates. Of course, most plants can be cultivated on a much wider variety of sites than they can successfully occupy under natural conditions, but such plantings strongly support the hypothesis that exclusion from sites dominated by other species is more likely the result of long-term competition than exacting requirements for a particular substrate.

A thesis emphasizing the subject of soil preference by bishop pine is in preparation by Mr. Ken Cole formerly of California State University, Los Angeles, and now at the University of Arizona. This manuscript was unavailable at the time of writing.

A safe generalization is that bishop pine tends to occur on sites which cannot be successfully occupied by other tree species, for these sites tend to be less favorable. For example, on Vandenberg AFB oak forest dominates the valley bottoms and slopes with what appear to be richer, finer textured soils, while the bishop pine is most abundant on steeper slopes and on the acid and almost certainly nutrient-poor sandstone-derived soils. This pattern

of conifers occupying less favorable sites than their broadleaf competitors is repeated in many places in the world, including the humid tropics (Whitmore 1975).

Seed release, seedling establishment, and early growth. Bishop pine is a closed-cone conifer, with cones described in botanical terms as "serotinous." This means that they do not open at maturity, but rather persist on the tree in a closed condition for many years. This trait, as discussed in another section, is found in quite a number (probably at least 20 species) of conifers in a variety of ecological settings around the world. The common feature of all closed-cone trees is that they experience crown fires at rather frequent intervals. Natural selection has apparently lead to the adoption of the serotinal behavior as a means of exploiting the habitat opened by fire.

Individuals in past times whose cones were not serotinal would, of course, have established seedlings without fire. But most of these seedlings would have been killed by fire. The trees that retained closed cones would then have relatively more seed available to repopulate the burned forest. So long as fires were frequent enough that the maturation of seedlings established at times other than after fires was unlikely, the evolutionary pressure must have favored the closed-cone habit.

The bishop pines on VAFB very definitely show the closed-cone character. As a result, the vast majority of trees became established after fires, and the stands tend to be even-aged. But although most cones will open only gradually and partially without fire, cones on dead branches will open if exposed to the sun. Others are attacked by squirrels, and a few seeds made available in this way. Bulldozing or cutting of trees generally results in some cones being broken and scattered about. As a result, it is not uncommon to observe seedlings and saplings of obviously younger age growing among older even-aged trees dating from a single past fire.

On VAFB this non-fire related establishment does not seem, on the whole, to be or have been of major importance. However, on Santa Cruz Island, past records as well as current observation make it clear that areas of bishop pine once badly degraded by sheep grazing are now dense pine forests without there having been any fires. This can only have happened if there had been substantial seedling establishment without fire. While it may be that the island forms of bishop pine are different in this respect from the mainland trees, it is reasonable to hypothesize that even on VAFB fire may not be absolutely necessary to maintain bishop pine stands.

The northern populations of bishop pine are reported not to have so marked a closed cone behavior. The pines on Santa Cruz Island seem also more likely to have open cones though nearly all trees would still definitely deserve to be called closed-cone trees. On VAFB, few open cones are to be seen except on dead branches. However, at least one tree in the Pine Canyon area was observed to have large numbers of open cones, including ones on relatively young branches. This tree was of the remorata type.

When fires do occur, the cones open as the result of the melting of the resin cementing the cone scales together. Observations by Mahrdt (personal communication) during the experimental burn of 1974 indicated that many, perhaps most, of the cones open at least part way during or immediately after fire. But winged seeds require some agitation to be

dislodged from the cone, and seed fall is probably protracted for weeks. In the 1974 burn, not all the cones were heated enough to cause them to open at the time of the fire, and some of these have gradually opened since, resulting in at least some seed dispersal evidently taking place two or more years after the fire.

The seeds of bishop pine are small and winged, and could, under the proper conditions, be blown considerable distances. The vast majority, however, fall directly beneath the trees. For example, no seedlings of bishop pine have been observed to have established in the mineral soil exposed by a very large fire break around the area of the 1974 burn. On the other hand, it has been noted (Metcalf 1921) that bishop pine in the northern part of its range is an aggressive invader of old fields.

Seedling establishment on VAFB is probably most likely in late winter and early spring. Observations on VAFB and reports in the literature seem to indicate that bishop pine seedlings, like those of many conifers probably do not establish well in dense litter, and this is undoubtedly another one of the reasons that seedling establishment in dense stands is low except after fire.

Early seedling survival has been estimated from observations in the 1974 burn. Seedling counts in 1975 and 1977 showed that only 38 of 110 seedlings survived, indicating mortality of about 60% in two years. The survival was, however, patchy.

Well established seedlings grow rapidly. In one area of the 1974 burn about 42% of the two and a half year old seedlings were 50 centimeters or over in height, with 107 centimeters being the largest height recorded.

Cone production begins early in bishop pine. A three year old individual with a single cone has been noted, but so far none of the seedlings in the 1974 burn have produced cones. Open-grown saplings about seven years old which dated from destruction of a portion of a stand by a bulldozer were noted to have many cones.

Later growth and maturity. The growth of trees in dense stands rapidly declines as the canopy closes. In some circumstances "dog's hair" stands form which are reminiscent of the dense areas of reproduction sometimes found in ponderosa pine. In such areas, expression of dominance seems to be gradual, and the trees will generally be reasonably straight. But in older and more open stands the trees develop flat-topped spreading crowns, and a central axis is lost. In exposed locations, the trees may be wind-shaped, with a pronounced lean. In other cases, where the trees grow as scattered individuals, the trees branch at or near the base to form bush-like individuals. This tendency for branching and spreading growth is at least in part hereditary, as the northern forms are much less prone to have poor form.

Bishop pine does not seem to be a very long-lived tree. The oldest tree found on VAFB was 72 years old. On Santa Cruz Island, which has apparently not had any major fires for many years, the older trees often show signs of senescence. In the Pine Canyon area of VAFB where the oldest stands are located (c. 60-70 years old), many dead trees are present. Others are alive but show evidence of heart-rot. It is doubtful that natural stands would be

healthy beyond an age of 100 years. The short life-span of bishop pine is not surprising since Monterey pine, its close relative, also is known to be short-lived.

Dead trees also quickly succumb to wood rotting organisms and termites. Trees dead less than three years in the 1974 burn are beginning to fall down as the result of termites and rot destroying the main roots.

Fire and mature trees. As has already been discussed, the natural pattern under which the bishop pine evolved must have included frequent destructive crown fires. Such fires are made more likely by the tendency for bishop pine to grow in dense stands. Such crown fires would almost inevitably kill nearly all trees. Bishop pine probably did not very often experience light ground fires under primeval conditions. If it had, it probably would not be a closed cone species, since light fires favor individuals that shed seed each year.

However, under current conditions, it does sometimes happen that fires will burn through a stand without producing a crown fire, as may be readily seen by the "cat-faces" (fire scars) on large trees. An example is a stand of trees just off Radio Receiver Road near the 1974 burn. It is also possible for fires to burn in a complicated pattern, leaving pockets of older trees in the midst of a younger stand. Such an area lies to the east of Lucio Road in South Vandenberg.

The southern form of bishop pine rarely occur in mixed stands with other trees. On VAFB an occasional oak may be found among the pines, especially where the pines occur in steep canyons. In a few instances toyon bushes large enough to be called trees were noted.

It seems that the bishop pine is a species adapted better to coexistence with shrubs than with other trees, and this accounts, in part, for its adaptations to crown fires. Nearly all stands on VAFB have at least some shrubs in them, and in many areas the pines occur as individuals scattered in the chaparral. The most common associates of the bishop pine are the interior live oak (Quercus wislezenii) and Pecho mountain manzanita (Arctostaphylos pechoensis var. viridissima).

Economic value. Bishop pine has been a commercial species in the northern part of its range, but use of the southern stands has probably only been local. At present there is no commercial use of the southern stands (Critchfield, personal communication). Aside from use as firewood and posts, it does not seem likely that any commercial use of natural stands will develop. However, in natural conditions bishop pine tends to occur on poor soils. When planted on better soils, the growth of bishop pine is often phenomenal, and it has attracted considerable attention as a possible forest plantation species, especially in New Zealand, Australia, and England (Fielding 1961, Shelbourne 1974, Everard and Fourt 1974).

So far, planting of bishop pine has been much less than that of its close relative, Monterey pine, which now is a major part of the forest industry in New Zealand, Australia, Chile, and other countries. But since the characteristics of bishop pine are much like those of Monterey pine, there is good reason to suppose that bishop pine might some day come to be of much greater importance than it is at present.

Evidence from existing plantations (Doran 1974, Shelbourne 1974) seems to suggest that on the best sites bishop pine is not quite so productive as Monterey pine, although it seems to be hardier, and is probably more productive on poorer soils and at higher elevations. It has also been found that the northern, or "blue" race, is generally better suited to cultivation (Doran 1974, Shelbourne 1974).

Considering the potential of bishop pine for use in forestry, it is important that natural stands be maintained as sources of genetic material for selection and breeding. Even though the southern races have not yet found favor as plantation species, they may be invaluable in breeding programs. This is especially likely if the goal of such a breeding program would be a tree adapted to less fertile or dryer sites.

D. B. Fourt (personal communication 1977) has suggested that the unusual chemical composition of the turpentines of the southern races may be potentially of value "in the production of essential oils, flavors, or scents."

Scientific and esthetic value. The extraordinary variability of bishop pine has already been discussed. This extreme variability and the presence of genetically isolated forms makes the species unique among the pines of the world, and therefore of considerable interest to tree geneticists and population geneticists. It is certain that much can be learned from a detailed study of the population genetics and geographic variation of bishop pine. Since such studies depend heavily upon the availability of stands across the entire range of the species, it becomes important to maintain as many of the populations as possible in a natural condition. This requires that the genetic makeup of the population not be disrupted by the interplanting of strains from other portions of the range (Critchfield 1977, personal communication).

Plant geographers and plant ecologists also find the species of interest, both as a representative of the closed-cone conifer forest that is once believed to have dominated coastal California, and as a modern day example of a tree species adapted to existence in what is (in the southern part of its range) predominately a chaparral and oak woodland environment.

Finally, there is the esthetic value of bishop pine. It is tempting to become poetic at this point and speak of "ancient denizens of the fog-bound coast" or "lichen-clad relics of the time when mammoths ruled in Lompoc." Such temptations will be bypassed and it will simply be stated that there is no doubt that many persons would ascribe a very great esthetic value to the bishop pine in its natural and undisturbed setting. Protection of the stands could be justified on these grounds alone.

DESCRIPTIONS OF THE STANDS

In this section a brief description of each of the stands sampled will be given. The stands are located on sections of the VAFB map in figures 1 and 2. The areas sampled and number of trees measured are shown in Table 1.

In the following discussions, the distribution of ages indicated for each of the stands is based on ring counts of increment cores. Bishop pine is a very plastic species, and it is able both to withstand considerable suppression, and also to grow very rapidly when released from suppression.

Interpreting age from ring counts in such a species is necessarily subject to error, especially for the trees showing extremes in growth rate. In many of the stands, suppression is extreme, and there may be a tendency to underestimate the age in some cores. Some of the range in age noted in the stands may arise from such difficulties. However, bishop pine can establish seedlings over a period of at least two years after a fire, as has been noted in the test burn on the base, and seedlings will also occasionally establish at other times. For this reason, it seems safest to stage stand ages as ranges rather than specific numbers. In the graphs in which stand features are plotted against age, the most likely predominant age has been selected.

The compass headings indicated for each transect are recorded as degrees clockwise from magnetic north.

Stand 1

Stand 1 is situated west of Santa Ynez Ridge Road approximately one-half mile south of Radio Receiver Road. It is located on a flat ridge top immediately adjacent to the roadway.

The trees are homogeneous in form and age. They are healthy, erect, single-stemmed individuals between 22-25 years of age. Evidence of severe squirrel predation and previous bark beetle predation was noted. The shrub cover is moderately dense.

Two 30 m transects were laid out end to end along 326°. They were each sampled to a width of 1 m on both sides.

Stand 2

Stand 2 is located one-half mile northwest of the Pine Canyon Gate. The sampling area is situated on the northwest facing slope of a small northeast oriented side canyon off Pine Canyon proper. The slopes of the smaller canyon are steep with an angle of 34°.

The stand on the northwest slope is open. The oldest trees of all the stands sampled are found in this stand. The ages are varied, ranging from 40-75 years, but the majority of trees are between 60-70 years old. The tree forms are quite variable as well. They may be large, erect and single-stemmed, or they may be knarled, twisted and highly branched. Most of the trees are infected with galls and have had most of the cones removed by squirrels. A large area approximately 100 m^2 within transect 2 was completely bare of cover except for the presence of 24 charred tree stumps. No other signs of previous fire were noted. The shrub cover is also sparse. Two herbs were noted in the transects, the bracken Pteridium aquilinum and deerweed, Lotus scoparius.

Two transects were sampled in this area. They were both aligned down the face of the slope. Transect 1 was oriented along 304° and Transect 2 was oriented along 282°. They were each 25 m long and were sampled to a width of 15 m on both sides to total an area of 1500 m^2.

Stand 3

Stand 3 is located on the southwest facing slope of Spring Canyon and is southeast of SLC #4. It is approximately 425 m from Transmitter 3 along the ridge top.

The stand is sparse but healthy. The trees are highly branched at the base and bear large numbers of cones. No animal predation was noted in this area. Two age classes are represented on the slope. The few live fire-scarred trees present are 40-47 years old, but the majority of trees are between 10-15 years old. None of the older trees occurred in the sample.

The shrub cover is dense. Two 30 m transects were aligned with the slope at an angle of 12° and oriented along 222°. They were each sampled to 15 m on both sides to total an area of 1800 m^2.

Stand 4

Stand 4 is located on a northeast facing slope 15 m from Stand 3 on the opposite side of the same ridge.

As with Stand 3, many indications of fire are present. There are two age classes. The older trees outside the area sampled are fire-scarred and are approximately 43 years old. The younger trees that are within the transect area are 10-15 years old. The character of Stand 4 is very different from Stand 3. The trees in this dense stand are smaller, thinner, and single-stemmed. This stand is one of the least vigorous sampled. Most of the trees exhibit yellowing leaves and many trees are recently dead.

The shrub cover is sparse. Two 15 m transects were oriented across the face of the slope at an angle of 5° and with an orientation of 330°. They were aligned parallel to each other and were each sampled on both sides to a width of 1 m. A total area of 60 m^2 was sampled.

Stand 5

Stand 5 is located 210 m northwest of the Arguello Blvd. and Santa Ynez Ridge Road intersection on a relatively flat ridge top.

The stand is very dense with thin single-stemmed trees which have their leaves restricted to a few upper branches. There is little cone production and the trees show signs of squirrel predation. Many charred cones litter the forest floor as do the remains of burned Ceonothus impressus. The stand is homogeneous in age and is between 13-15 years old. The shrub cover is moderately dense.

One 30 m transect oriented along 241° was sampled to a width of 1 m on each side. The total sample area being 60 m^2.

Stand 6

Stand 6 is located along the east side of Lucio Road 100 m from the road's intersection with Arguello Blvd. The forest floor slopes toward

the east at an angle of 10° for 40 m, and then slopes upward again at the same angle to form a shallow canyon.

The stand consists of large trees in moderate health. Although the canopy cover is dense, the forest is fairly open. Cone predation by squirrels is severe. The trees are uniformly 30-35 years old. Only one charred tree stump was noted as an indication of previous fire within the stand.

The shrub cover was sparse and clumped to moderately dense. Three 30 m transects were sampled. The first was oriented along the canyon bottom at 335°. The other two transects were laid down approximately parallel to each other down along the west facing slope. Transect 2 was oriented along 291° while Transect 3 was oriented along 286°. They were each sampled to 1 m on both sides to total an area of 180 m^2.

Stand 7

Stand 7 is located 40 m northwest of the road end parking area off Lucio Road. It is situated on a relatively flat ridge top that slopes gently to the northwest at an angle of 5°.

The stand is moderately dense and is similar in character to Stand 6. The trees are large and the forest floor is open. There is evidence of heavy cone predation and moderate bark predation by squirrels. There are a number of dead saplings remaining erect.

A few charred cones are present in the leaf litter. The trees are uniformly 18-20 years old.

The shrub cover is very sparse but there are a number of scattered herb species. There are also a few live seedlings of Q. wislizenii. There was an abundance of Rhus (Toxicodendron) diversiloba in the stand but none occurred within the transects. The herb species present within the transects were Satureja douglasii, Salvia spathacea, and Galuim sp. Three 30 m transects were sampled in the approximate direction of the slope. Transect 1 was oriented at 310°. Transect 2 at 302°, and Transect 3 at 298°. Each transect was sampled to 1 m widths on both sides to total a sampled area of 180 m^2.

Stand 8

Stand 8 is located approximately 300 m northwest of the Lucio Road parking area. The stand is situated on a southwest facing slope at an angle of 30°. The trees of this dense stand are uniformly 19-20 years old. They are vigorous, single-stemmed individuals, but many of them have been infected by cankers. There is very little evidence of cone or bark predation by squirrels in this stand. A few charred cones present in the leaf litter are the only indications of previous fire.

The shrub cover is low and fairly dense.

Two transects were sampled across the face of the slope. They were parallel to each other and were oriented along 118°. The first transect

was 20 m long, and the second, located 5 m further up the slope from the first transect, was 10 m long. Both sides of the transects were sampled to 1 m widths with the total sampled area being 60 m^2.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

NONLINEAR GUIDANCE FOR AIR-TO-AIR MISSILES

Prepared by:	Jan F. Andrus, PhD
Academic Rank:	Associate Professor
Department and University:	Department of Mathematics University of New Orleans

Assignment:

(Air Force Base)	Eglin AFB FL
(Laboratory)	Air Force Armament Laboratory
(Division)	Guided Weapons Division
(Branch)	Systems Analysis & Simulation Branch

USAF Research Colleague:	Mario J. Caluda, PhD
Date:	August 5, 1977
Contract No:	F44620-75-C-0031

Nonlinear Guidance for Air-to-Air Missiles

by

J.F. Andrus

ABSTRACT

The problem is that of on-board, real-time guidance of an air-to-air missile engaging an accelerating target vehicle. It is assumed that the motion of the target can be predicted. Optimal guidance methods are developed starting with nonlinear translational equations of motion. The performance index to be minimized is the sum of terms representing the square of the miss-distance and the integral of the sum of the squares of the rates of change of angles defining the direction of the instantaneous velocity vector of the missile.

First a guidance method, based upon a series expansion of the control as a function of time, is developed for the case of three dimensions and variable V_M, where V_M is the magnitude of the velocity of the missile. It is assumed that V_M is known as a function of time.

Then a second optimal guidance method is developed which is based upon a closed-form solution to the two-dimensional translational equations for the case of constant V_M. The guidance command on each guidance cycle is obtained from the simultaneous solution to two equations which involve elliptic integrals.

ACKNOWLEDGEMENTS

The author is particularly indebted to Dr Mario J. Caluda for his guidance in determining the direction in which the research project should progress. The author is also thankful for Major Jim Anderson's support including his explanations of principal physical factors involved in missile engagements. The very helpful assistance of Mr Fred O'Brien of Auburn University is also greatly appreciated.

INTRODUCTION

The rapid development of digital computer technology and improvements in estimation methods now make it possible to incorporate sophisticated real-time guidance algorithms into the control systems of air-to-air missiles. This report develops nonlinear optimal guidance algorithms for two- and three-dimensional encounters. The development is based upon the assumption that the time-history of the target can be estimated reasonably accurately. The new guidance methods are expected to be more accurate than proportional navigation which is non-optimal for the cases of accelerating targets and nonlinear systems.

THE MODEL

The coordinate system to be employed is shown in Figure 1 where the angles, α_A and α_E, define the direction of the instantaneous missile velocity vector.

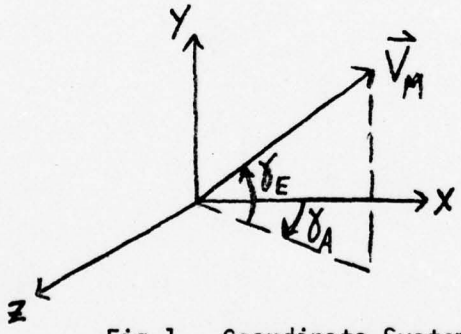


Fig 1. Coordinate System

The equations of relative motion of the target and missile are

$$\ddot{\vec{R}} = \vec{V}_T(t) - V_M(t)\vec{p}(\vec{\Gamma})$$

where

$$\vec{\Gamma} = \begin{bmatrix} \gamma_A \\ \gamma_E \end{bmatrix} \quad , \quad \vec{p} = \begin{bmatrix} \cos\gamma_E \cos\gamma_A \\ \sin\gamma_E \\ \cos\gamma_E \sin\gamma_A \end{bmatrix}$$

and $\vec{V}_T(t)$ is the predicted time-history of the velocity of the target (the predictions may change between guidance cycles). Likewise $V_M(t)$ is the predicted time-history of the velocity magnitude of the missile, the vector $\vec{R}$ is the difference between the positions of the target and the missile. It defines the line-of-sight. It is assumed that the initial values, $\vec{R}_o$ and $\vec{r}_o$, are known.

The control will be $\vec{u}=\dot{\vec{r}}$ and the performance index to be minimized is

$$J = \tfrac{1}{2}k|\vec{R}(T)|^2 + \tfrac{1}{2}\int_0^T |\vec{u}|^2 dt$$

where T is time-to-go. The index includes the square of the miss-distance andan integral of the square of the magnitude of the control vector. The integral is included as a penalty term in order to prevent unreasonably large control commands which would be impossible to implement or which would lead to large losses in the energy of the missile.

CONDITIONS OF OPTIMALITY

In summary, the control problem is defined as follows:

$$\dot{\vec{R}} = \vec{V}_T - V_M\vec{p}(\vec{r})$$

$$\dot{\vec{r}} = \vec{u}$$

$$\dot{\alpha} = \tfrac{1}{2}\vec{u}^T\vec{u}$$

$$\underset{\vec{u},T}{\text{Min}}\ J = \tfrac{1}{2}k\vec{R}^T(T)\vec{R}(T) + \alpha(T)$$

The necessary conditions of optimality will now be derived using the Minimum Principle.

The Hamiltonian is:

$$H = \vec{\lambda}^T(\vec{V}_T - V_M\vec{p}) + \vec{\beta}^T\vec{u} + \tfrac{1}{2}\phi\vec{u}^T\vec{u}$$

where $\vec{\lambda}$, $\vec{\beta}$, and ϕ are the costate variables.

The costate equations are:

$$\dot{\vec{\lambda}} = -(\partial H/\partial \vec{R})^T, \dot{\vec{\beta}} = -(\partial H/\partial \vec{\Gamma})^T, \dot{\phi} = -\partial H/\partial \alpha$$

For the particular problem the equations are:

$$\dot{\vec{\lambda}} = \vec{0}, \dot{\vec{\beta}} = V_M A^T \vec{\lambda}, \dot{\phi} = 0$$

where $A = \partial \vec{p}/\partial \vec{\Gamma}$.

The transversality conditions are

$$\vec{\lambda}(T) = (\partial J/\partial \vec{R}(T))^T, \quad \vec{\beta}(T) = (\partial J/\partial \vec{\Gamma}(T))^T,$$
$$\phi(T) = \partial J/\partial \alpha(T), \quad H(T) = 0 .$$

For the particular problem the latter conditions are:

$$\vec{\lambda}(T) = k\vec{R}(T), \quad \vec{\beta}(T) = \vec{0}, \quad \phi(T) = 1, \quad H(T) = 0 .$$

The costate and transversality equations imply

$$\vec{\lambda}(t) \equiv k\vec{R}_F$$
$$\phi(t) \equiv 1$$
$$k(\vec{V}_{TF} - V_M \vec{p}_F)\vec{R}_F + \tfrac{1}{2}|\vec{u}_F|^2 = 0 \qquad (1)$$

where the subscript F indicates evaluation at the final time T.

The condition $\partial H/\partial \vec{u} = \vec{0}^T$ defines the optimum control . This condition yields $\vec{\beta} + \phi\vec{u} = 0$; i.e.

$$\vec{u} \equiv -\vec{\beta}$$

Since $\vec{\beta}_F = \vec{0}$, $\vec{u}_F = \vec{0}$ so that equation (1) gives

$$\vec{R}_F \cdot \dot{\vec{R}}_F = 0$$

The physical interpretation of the latter condition is that, at the final point of the optimum trajectory, the relative position and velocity vectors are orthogonal. This situation is illustrated in Figure 2. If V_M is several times larger than $|\dot{\vec{V}}_{TF}|$, then the final velocity of the missile will be nearly orthogonal to the line-of-sight.

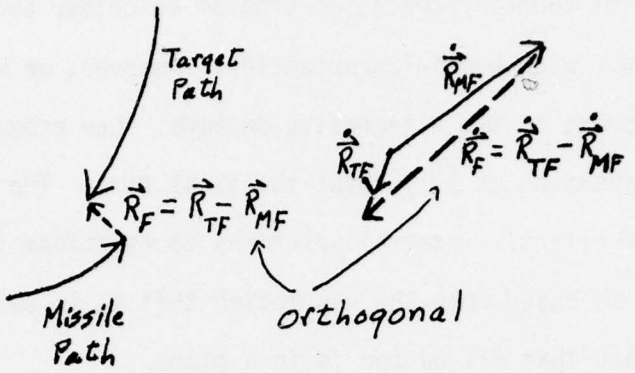


Figure 2. Final Relative Positions and Velocities

In order to eliminate $\vec{\beta}$ from the problem, one may differentiate both sides of the equation $\dot{\vec{\Gamma}} = -\vec{\beta}$ with respect to time and then employ the costate equation $\dot{\vec{\beta}} = k V_M A^T \vec{R}_F$ to obtain $\ddot{\vec{\Gamma}} = -k V_M A^T \vec{R}_F$.

In summary, the necessary conditions of optimality are:

$$\dot{\vec{R}} = \vec{V}_T(t) - V_M(t)\,\vec{p}(\vec{\Gamma}) \tag{2}$$

$$\ddot{\vec{\Gamma}} = -k V_M(t) A^T \vec{R}_F \tag{3}$$

with boundary conditions

$$\vec{R}(0) = \vec{R}_0 , \quad \vec{\Gamma}(0) = \vec{\Gamma}_0 , \quad \dot{\vec{\Gamma}}(T) = \vec{0} , \quad \vec{R}_F \cdot \dot{\vec{R}}_F = 0 .$$

Recall that A is a matrix whose elements depend upon $\vec{\Gamma}$.

The above conditions give a well-defined two-point boundary-condition problem which consists of determining values of T, $\dot{\vec{\Gamma}}(0)$, and $\vec{R}_F$ which will satisfy the equations $\vec{R}_F \cdot \dot{\vec{R}}_F$, $\dot{\vec{\Gamma}}_F = \vec{0}$, and $\vec{R}_F = \int_0^T [\vec{V}_T - V_M \vec{p}(\vec{\Gamma})]\,dt$ with $\vec{\Gamma}$ being the solution to equation (3). There are six unknowns and

six equations to be satisfied.

The two-point boundary-condition problem described above can be solved numerically; e.g., with quasi-linearization. However, we will proceed with the development of two alternative methods. One procedure will be based upon a series expansion of $\vec{r}(t)$ about the final time. The other method will make use of elliptic integral solutions to equations (2) and (3). The latter solution is based upon the assumption that V_M is constant on each guidance cycle and that all motion is in a plane.

AN APPROXIMATE SOLUTION

In order to obtain an approximation for $\vec{r}(t)$, it will be expanded in a Taylor series through third-order terms about t=T. Additional terms can be taken as necessary; for example, a particularly simple expansion can be obtained through fifth-order terms for the case of V_M constant. The expansion is about the final time because the condition $\dot{\vec{r}}(T)=\vec{0}$ leads to simplifications. Employing equations (3) and the condition $\dot{\vec{r}}(T)=\vec{0}$, it is easily shown that

$$\ddot{\vec{r}}_F = -k V_{MF} A_F^T \vec{R}_F \;, \quad \dddot{\vec{r}}_F = -k \dot{V}_{MF} A_F^T \vec{R}_F \;.$$

Assuming the remainder terms to be relatively small, we have the approximation

$$\vec{r}(t) \doteq \vec{r}_F + (T-t)\dot{\vec{r}}_F + \frac{1}{2!}(T-t)^2 \ddot{\vec{r}}_F + \frac{1}{3!}(T-t)^3 \dddot{\vec{r}}_F$$

so that

$$\vec{r}(t) \doteq \vec{r}_F - \frac{1}{2} k V_{MF} (T-t)^2 A_F^T \vec{R}_F - \frac{1}{6} k \dot{V}_{MF} (T-t)^3 A_F^T \vec{R}_F \qquad (4)$$

For the case of $\dot{V}_{MF}=0$, the formula becomes especially simple.

Similarly $\vec{p}(\vec{r})$ may be expanded. The time derivatives will first be expressed:

$$\frac{d}{dt}\vec{p}\,\Big|_T = A_F \dot{\vec{r}}_F = 0$$

$$\frac{d^2}{dt^2}\vec{p}\,\Big|_T = A_F \ddot{\vec{r}}_F = -k V_{MF} A_F A_F^T \vec{R}_F$$

$$\left.\frac{d^3\vec{p}}{dt^3}\right|_T = A_F \dddot{\vec{\Gamma}}_F = -k\dot{V}_{MF} A_F A_F^T \vec{R}_F$$

Therefore,

$$\vec{p}(\vec{\Gamma}) = \vec{p}_F - \tfrac{1}{2} k V_{MF} (T-t)^2 A_F A_F^T \vec{R}_F - \tfrac{1}{6} k \dot{V}_{MF} (T-t)^3 A_F A_F^T \vec{R}_F$$

Equations (2) may now be integrated as follows:

$$\vec{R}_F = \vec{R}_o + \int_o^T \vec{V}_T\, dt - \left(\int_o^T V_M\, dt\right)\vec{p}_F + \tfrac{1}{2} k V_{MF} \left(\int_o^T (T-t)^2 V_M\, dt\right) A_F A_F^T \vec{R}_F + \tfrac{1}{6} k \dot{V}_{MF} \left(\int_o^T (T-t)^3 V_M\, dt\right) A_F A_F^T \vec{R}_F \quad (5)$$

Setting t=0 in equations (4) yields

$$\vec{\Gamma}_F - \vec{\Gamma}_o = \tfrac{1}{2} k \left(V_{MF} + \tfrac{1}{3}\dot{V}_{MF} T\right) T^2 A_F^T \vec{R}_F \quad (6)$$

The unknown vector $A_F^T \vec{R}_F$ will be eliminated between equations (5) and (6). Solving equations (6) for $\tfrac{1}{2} k A_F^T \vec{R}_F$ and substituting into equations (5) gives

$$\vec{R}_F = \vec{R}_o + \int_o^T \vec{V}_T\, dt - \left(\int_o^T V_M\, dt\right)\vec{p}_F + \frac{V_{MF}\int_o^T (T-t)^2 V_M\, dt + \tfrac{1}{3}\dot{V}_{MF}\int_o^T (T-t)^3 V_M\, dt}{\left(V_{MF} + \tfrac{1}{3}\dot{V}_{MF} T\right) T^2} A_F (\vec{\Gamma}_F - \vec{\Gamma}_o) \quad (7)$$

Substitution of the right-hand members of equations (7) into equations (6) gives

$$\vec{\Gamma}_F - \vec{\Gamma}_o = \tfrac{1}{2} k \left(V_{MF} + \tfrac{1}{3}\dot{V}_{MF} T\right) T^2 A_F^T \left[\vec{R}_o + \int_o^T \vec{V}_T\, dt - \left(\int_o^T V_M\, dt\right)\vec{p}_F\right] + \tfrac{1}{2} k \left[V_{MF}\int_o^T (T-t)^2 V_M\, dt + \tfrac{1}{3}\dot{V}_{MF}\int_o^T (T-t)^3 V_M\, dt\right] A_F^T A_F (\vec{\Gamma}_F - \vec{\Gamma}_o) \quad (8)$$

Equations (8) are two equations in the two unknown components of $\vec{\Gamma}_F$. All other quantities are theoretically known if we assume T is computed numerically in some "outside loop" which minimizes miss-distance with respect to T assuming $\vec{\Gamma}_F$ depends upon T through equations (8). (It is also possible to solve for $\vec{\Gamma}_F$ and T by simultaneous numerical solution of equations (8) and the condition $\vec{R}_F \cdot \dot{\vec{R}}_F = 0$ with $\vec{R}_F$ given by equations (7) and $\dot{\vec{R}}_F$ given by equations (2).)

Once $\vec{\Gamma}_F$ and T have been computed, the control $\vec{u}=\dot{\vec{\Gamma}}$ can be calculated from the following equations obtained by differentiating both members of equations (4):

$$\vec{u}(t) = k(T-t)\left[V_{MF} + \tfrac{1}{2}\dot{V}_{MF}(T-t)\right] A_F^T \vec{R}_F$$

with $\vec{R}_F$ coming from equations (7).

The problem of obtaining a guidance command is essentially that of numerically solving the two nonlinear equations (8) for γ_{AF} and γ_{EF}. Integrals, such as $\int (T-t)^3 V_M dt$ can be easily computed if $V_M(t)$ is approximated by means of an elementary function. In order to simplify equations (8), it is observed that

$$A^T = \begin{bmatrix} -\cos\gamma_E \sin\gamma_A & 0 & \cos\gamma_E \cos\gamma_A \\ -\sin\gamma_E \cos\gamma_A & \cos\gamma_E & -\sin\gamma_E \sin\gamma_A \end{bmatrix}$$

$$A^T A = \begin{bmatrix} \cos^2\gamma_E & 0 \\ 0 & \sin^2\gamma_E \end{bmatrix} = D(\gamma_E), \quad A^T \vec{\rho} = \vec{0}$$

Therefore, equations (8) may be written as

$$\vec{\Gamma}_F - \vec{\Gamma}_0 = A_F^T \vec{a} + b\, D_F (\vec{\Gamma}_F - \vec{\Gamma}_0)$$

where

$$\vec{a} = \tfrac{1}{2} k (V_{MF} + \tfrac{1}{3}\dot{V}_{MF} T) T^2 (\vec{R}_0 + \int_0^T \vec{V}_T dt)$$

$$b = \tfrac{1}{2} k \left[V_{MF} \int_0^T (T-t)^2 V_M dt + \tfrac{1}{3}\dot{V}_{MF} \int_0^T (T-\tau)^3 V_M dt\right]$$

Letting $\eta = \tan^{-1}(a_3/a_1)$, the equations can be written in terms of components as follows:

$$(\gamma_{AF} - \gamma_{A0})(b\cos^2\gamma_{EF} - 1) = \sqrt{a_1^2 + a_3^2}\sin(\gamma_{AF} - \eta)\cos\gamma_{EF} \qquad (9\text{-}A)$$

$$(\gamma_{EF} - \gamma_{E0})(b\sin^2\gamma_{EF} - 1) = \sqrt{a_1^2 + a_3^2}\cos(\gamma_{AF} - \eta)\sin\gamma_{EF} - a_2\cos\gamma_{EF} \qquad (9\text{-}B)$$

If one wishes to decouple equations (9), it is possible to solve equation (9-B) for γ_{AF} in terms of γ_{EF} and then to substitute into equation (9-A) in order to obtain an equation involving γ_{EF} alone.

In the planar case, we may take $\gamma_E \equiv 0$ and $a_2 = 0$, in which case equation (9-B) is trivially satisfied and equation (9-A) reduces to the equation

$$(b-1)(\gamma_{AF} - \gamma_{A0}) = \sqrt{a_1^2 + a_3^2}\,\sin(\gamma_{AF} - \eta)$$

which can be solved for γ_{AF} by numerically calculating the intersection of a line and a sine curve.

AN ELLIPTIC INTEGRAL SOLUTION

This section is concerned with expressing the solution to the two-point boundary-condition problem, defined by equations (2) and (3), in terms of elliptic integrals. It is assumed that, on each guidance cycle, motion is in the plane defined by $\gamma_E \equiv 0$, that V_M is constant, and that $\dot{\gamma}_A$ does not change sign.

In two dimensions equations (3) become

$$\ddot{\gamma}_A = k V_M (R_{1F} \sin\gamma_A - R_{3F} \cos\gamma_A)$$

Let $z = \dot{\gamma}_A$ so that $\ddot{\gamma}_A = z(dz/d\gamma_A)$. Then

$$z \frac{dz}{d\gamma_A} = k V_M (R_{1F} \sin\gamma_A - R_{3F} \cos\gamma_A)$$

Multiplying both sides by $d\gamma_A$ and integrating, we obtain

$$\tfrac{1}{2} z^2 = -k V_M (R_{1F} \cos\gamma_A + R_{3F} \sin\gamma_A) + C$$

Since $z(T) = \dot{\gamma}_A(T) = 0$, the constant is

$$C = k V_M (R_{1F} \cos\gamma_{AF} + R_{3F} \sin\gamma_{AF})$$

Recalling that $\dot{\gamma}_A = z$, we obtain

$$\dot{\gamma}_A = \pm \sqrt{2 k V_M \left[(R_{1F} \cos\gamma_{AF} + R_{3F} \sin\gamma_{AF}) - (R_{1F} \cos\gamma_A + R_{3F} \sin\gamma_A) \right]} \qquad (10)$$

The latter equation is based upon an assumption that $\dot{\gamma}_A$ always has the same sign on particular guidance cycle. This assumption is backed up, for the case of V_M constant, by the formula $\vec{\Gamma} \doteq k(T-t)A_F^T \vec{R}_F$ obtained in the preceeding section.

Equation (10) can be rewritten as follows:

$$\dot{\gamma}_A = \pm\sqrt{2kV_M R_F}\sqrt{\cos(\gamma_{AF}-\nu)-\cos(\gamma_A-\nu)} \tag{11}$$

where

$$\nu = \tan^{-1}(R_{3F}/R_{1F})$$

Observe that it is necessary that $\cos(\gamma_{AF}-\nu) \geq \cos(\gamma-\nu)$.

After separation of variables, equation (11) can be integrated to obtain

$$\pm T\sqrt{2kV_M R_F} = I_1(\gamma_{AF},\nu) \tag{12}$$

where

$$I_1(\gamma_{AF},\nu) = \int_{\gamma_{A0}}^{\gamma_{AF}} \frac{d\gamma_A}{\sqrt{\cos(\gamma_{AF}-\nu)-\cos(\gamma_A-\nu)}}$$

In two dimensions, equations (2) reduce to

$$\dot{R}_1 = V_{T1} - V_M\cos\gamma_A \tag{13-A}$$

$$\dot{R}_3 = V_{T3} - V_M\sin\gamma_A \tag{13-B}$$

Integrating equation (13-A), for example, we obtain

$$R_{1F} = R_{10} + \int_0^T V_{T1}\,dt - V_M\int_0^T \cos\gamma_A\,dt$$

$$R_{1F} = R_{10} + \int_0^T V_{T1}\,dt - V_M\int_{\gamma_{A0}}^{\gamma_{AF}} \frac{\cos\gamma_A}{\dot{\gamma}_A}\,d\gamma_A$$

$$R_{1F} = R_{10} + \int_0^T V_{T1}\,dt \mp \sqrt{\frac{V_M}{2kR_F}}\,I_c(\gamma_{AF},\nu) \tag{14}$$

where

$$I_c(\gamma_{AF},\nu) = \int_{\gamma_{A0}}^{\gamma_{AF}} \frac{\cos\gamma_A}{\sqrt{\cos(\gamma_{AF}-\nu)-\cos(\gamma_A-\nu)}}\,d\gamma_A$$

Similarly

$$R_{3F} = R_{30} + \int_0^T V_{T3}\,dt \mp \sqrt{\frac{V_M}{2kR_F}}\,I_s(\gamma_{AF},\nu) \tag{15}$$

where

$$I_s(\gamma_{AF},\nu) = \int_{\gamma_{A0}}^{\gamma_{AF}} \frac{\sin\gamma_A}{\sqrt{\cos(\gamma_{AF}-\nu)-\cos(\gamma_A-\nu)}}\,d\gamma_A$$

Substitution of the right-hand members of equation (2) into the transversality condition,

$\vec{R}_F \cdot \dot{\vec{R}}_F = 0$ gives

$$\vec{R}_F \cdot \vec{V}_{TF} - V_M (R_{1F} \cos\gamma_{AF} + R_{3F} \sin\gamma_{AF}) = 0$$

Therefore

$$\vec{R}_F \cdot \vec{V}_{TF} = V_M R_F \cos(\gamma_{AF} - \nu) \tag{16}$$

In summary equations (12), (14), (15), and (16) are four nonlinear equations in the unknowns γ_{AF}, R_{1F}, R_{3F}, and T. The solution to these equations gives a solution to the optimal control problem. It will be shown later that the integrals I_1, I_c, and I_s can be expressed in terms of elliptic integrals. Next we will give a preliminary analysis of these integrals and then proceed to analyze equations (12), (14), (15), and (16).

Let $\eta = \gamma_A - \nu$. Then $\eta_0 = \gamma_{A0} - \nu$ and $\eta_F = \gamma_{AF} - \nu$.

Hence

$$I_1 = \int_{\eta_0}^{\eta_F} \frac{d\eta}{\sqrt{\cos\eta_F - \cos\eta}}$$

$$I_c = \int_{\eta_0}^{\eta_F} \frac{\cos(\eta+\nu)}{\sqrt{\cos\eta_F - \cos\eta}}\, d\eta$$

$$I_s = \int_{\eta_0}^{\eta_F} \frac{\sin(\eta+\nu)}{\sqrt{\cos\eta_F - \cos\eta}}\, d\eta$$

Therefore

$$I_c = (\cos\nu)\, J_c(\gamma_{AF}, \nu) - (\sin\nu)\, J_s(\gamma_{AF}, \nu)$$

$$I_s = (\sin\nu)\, J_c(\gamma_{AF}, \nu) + (\cos\nu)\, J_s(\gamma_{AF}, \nu)$$

AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

3 OF 6
AD
A051624

where

$$J_c = \int_{\eta_0}^{\eta_F} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\, d\eta \quad , \quad J_s = \int_{\eta_0}^{\eta_F} \frac{\sin\eta}{\sqrt{\cos\eta_F - \cos\eta}}\, d\eta$$

The integral J_s may be integrated to obtain

$$J_s = 2\sqrt{\cos\eta_F - \cos\eta}\,\Big|_{\eta_0}^{\eta_F}$$

$$J_s = -2\sqrt{\cos\eta_F - \cos\eta_0}$$

Therefore, the evaluation of all of the integrals reduces to the evaluation of the two integrals, I_1 and J_c. Integrals I_1 and J_c can be expressed in terms of elliptic integrals.

Now equations (12), (14), (15), and (16) will be simplified and expressed in terms of the unknowns η_0, η_F, $|\vec{R}_F|$, and T.

Equations (14) and (15) may be written as

$$|\vec{R}_F|\cos\nu = R_1^* \mp \sqrt{\frac{V_M}{2kR_F}}\left(J_c\cos\nu - J_s\sin\nu\right) \qquad (17\text{-}A)$$

$$|\vec{R}_F|\sin\nu = R_3^* \mp \sqrt{\frac{V_M}{2kR_F}}\left(J_c\sin\nu + J_s\cos\nu\right) \qquad (17\text{-}B)$$

where $R_i^* = R_{i0} + \int_0^T V_{Ti}\, dt$. Multiplying equation (17-A) by $\cos\nu$ and adding it to equation (17-B) multipled by $\sin\nu$ gives:

$$|\vec{R}_F| = R_1^*\cos(\gamma_{A0} - \eta_0) + R_3^*\sin(\gamma_{A0} - \eta_0) \mp \sqrt{\frac{V_M}{2kR_F}}\, J_c$$

which may be written as

$$|\vec{R}_F| = (R_1^*\cos\gamma_{A0} + R_3^*\sin\gamma_{A0})\cos\eta_0 + (R_1^*\sin\gamma_{A0} - R_3^*\cos\gamma_{A0})\sin\eta_0 \mp \sqrt{\frac{V_M}{2kR_F}}\, J_c \qquad (18)$$

Now multiply equation (17-A) by $\sin\nu$ and subtract equation (17-B) multiplied by $\cos\nu$. This yields

$$R_1^*\sin(\gamma_{A0} - \eta_0) - R_3^*\cos(\gamma_{A0} - \eta_0) \pm \sqrt{\frac{V_M}{2kR_F}}\, J_s = 0$$

$$(R_1^*\sin\gamma_{A0} - R_3^*\cos\gamma_{A0})\cos\eta_0 - (R_1^*\cos\gamma_{A0} + R_3^*\sin\gamma_{A0})\sin\eta_0 \pm \sqrt{\frac{V_M}{2kR_F}}\, J_s = 0 \qquad (19)$$

Equations (17) will be replaced by the equivalent equations, (18), and (19).

Equation (16) may be written as

$$\left[V_{T1F}\cos(\gamma_{AO}+\eta_o)+V_{T3F}\sin(\gamma_{AO}+\eta_o)\right]=V_M\cos\eta_F$$

$$\left[(V_{T1F}\cos\gamma_{AO}+V_{T3F}\sin\gamma_{AO})\cos\eta_o - (V_{T1F}\sin\gamma_{AO}-V_{T3F}\cos\gamma_{AO})\sin\eta_o\right]=V_M\cos\eta_F \quad (20)$$

Equations (18), (19), (20) may now be written as

$$|\vec{R}_F|=|\vec{R}^*|\left[\cos(\varsigma^*-\gamma_{AO})\cos\eta_o-\sin(\varsigma^*-\gamma_{AO})\sin\eta_o\right]\mp\sqrt{\frac{V_M}{2kR_F}}J_C \quad (21)$$

$$|\vec{R}^*|\left[\sin(\varsigma^*-\gamma_{AO})\cos\eta_o+\cos(\varsigma^*-\gamma_{AO})\sin\eta_o\right]\mp\sqrt{\frac{V_M}{2kR_F}}J_S=0 \quad (22)$$

$$|\vec{V}_{TF}|\cdot\left[\cos(\varsigma-\gamma_{AO})\cos\eta_o+\sin(\varsigma-\gamma_{AO})\sin\eta_o\right]=V_M\cos\eta_F \quad (23)$$

where $\varsigma^*=\tan^{-1}R_3^*/R_1^*$ and $\varsigma=\tan^{-1}V_{T3F}/V_{T1F}$.

Let $\lambda^*=\varsigma^*-\gamma_{AO}$ and $\lambda=\varsigma-\gamma_{AO}$. Then λ^* and λ are the angles measured positive counter-clockwise from the initial velocity vector to the vectors $\vec{R}^*$ and $\vec{V}_{TF}^*$, respectively. Then equations (21), (22), (23), and (12) may be written as

$$|\vec{R}_F|=|\vec{R}^*|(\cos\lambda^*\cos\eta_o-\sin\lambda^*\sin\eta_o)\mp\sqrt{\frac{V_M}{2kR_F}}\int_{\eta_o}^{\eta_F}\frac{\cos\eta}{\sqrt{\cos\eta_F-\cos\eta}}\,d\eta \quad (24)$$

$$|\vec{R}^*|(\sin\lambda^*\cos\eta_o+\cos\lambda^*\sin\eta_o)=\mp\sqrt{\frac{2V_M}{RR_F}}\sqrt{\cos\eta_F-\cos\eta_o} \quad (25)$$

$$|\vec{V}_{TF}|\cdot(\cos\lambda\cos\eta_o+\sin\lambda\sin\eta_o)=V_M\cos\eta_F \quad (26)$$

$$\pm T\sqrt{2kV_MR_F}=\int_{\eta_o}^{\eta_F}\frac{d\eta}{\sqrt{\cos\eta_F-\cos\eta}} \quad (27)$$

Equations (24)-(27) are four nonlinear equations in the four unknowns, η_0, η_F, $R_F = |\vec{R}_F|$, and T. There are many ways in which one can attempt to solve these equations numerically. One such method will be discussed below. We first observe that λ and λ^* both depend upon T, and that the + and - signs of the $\pm$ symbol correspond to the cases $\dot{\gamma}_A > 0$ and $\dot{\gamma}_A < 0$, respectively.

The integrals, J_c and I_1 occurring in equations (24) and (27) are analyzed in the appendix, where they are expressed in terms of elliptic integrals. There it is assumed that the magnitude of the final velocity of the missile is several times that of the target and that γ_A varies by no more than $\pm \pi/2$ radians. It is shown that, without loss of generality, the problem may be restricted to the case of η decreasing with time. Then $\eta_F \doteq \pi/2$ and $\eta_0 \in [\eta_F, \pi]$ In this case, the minus sign of the $\pm$ symbol applies in equations (24), (25), and (27).

Equations (24)-(27) may be written as

$$|\vec{R}_F| = |\vec{R}^*| \cos(\eta_0 + \lambda^*) \mp \sqrt{\frac{V_M}{2k R_F}}\, J_c \tag{28}$$

$$|\vec{R}^*| \sin(\eta_0 + \lambda^*) = \pm \sqrt{\frac{V_M}{2k R_F}}\, J_s \tag{29}$$

$$|\vec{V}_{TF}| \cos(\eta_0 - \lambda) = V_M \cos\eta_F \tag{30}$$

$$\pm T\sqrt{2kV_M R_F} = I_1 \tag{31}$$

Equations (29) and (31) combine to give

$$\frac{|\vec{R}^*|}{TV_M} \sin(\eta_0 + \lambda^*) = \frac{J_s}{I_1} \tag{29'}$$

Assuming T is determined independently in some "outside loop", equation (29)'may be solved numerically for η_0 with $\cos\eta_F$ given in terms of η_0 by means of equation (30). Then $|\vec{R}_F|$ may be computed by means of the equation

$$|\vec{R}_F| = |\vec{R}^*| \cos(\eta_0 + \lambda^*) - T V_M \frac{J_c}{I_1}$$

At this point, we note a certain peculiarity that can occur when the latter computations give $|\vec{R}_F| < 0$. This situation indicates that the substitution $R_{1F} = |\vec{R}_F| \cos\nu$ and $R_{3F} = |\vec{R}_F| \sin\nu$ which were effected during the derivation of the elliptic integral solutions, should be interpreted as $R_{1F} = (-|\vec{R}_F|)\cos\nu^*$ and $R_{3F} = (-|\vec{R}_F|)\sin\nu^*$, where $\nu^* = \nu + 180^\circ$. In this case the solution gives $-|\vec{R}_F|$ and ν^* rather than $|\vec{R}_F|$ and ν.

CONCLUSIONS AND RECOMMENDATIONS

It has been shown that it is possible to develop simple optimal guidance algorithms for air-to-air missiles starting with nonlinear translational equations of motion. Past efforts have concentrated upon linearized equations of motion.

The methods developed herein have received very little application. Before they can be fully appreciated, it is necessary that they be applied to a wide range of possible encounters and the results compared to those of other guidance methods. Moreover, the numerical techniques employed in the guidance methods need to be further refined and systematized.

In the case of the method based upon elliptic integrals, there are a number of extensions which should be attempted. The additional considerations include the following:

--introduction of variable V_M using perturbation methods

--extension into three dimensions

--allow variations in the control angles of more than $\pm \pi/2$ radians

--consideration of the possibility that the control angles may not always vary monotonically with time

--addition of an improved method for calculating time-to-go

--simplification of the equations by expanding η_F about $\pi/2$ (if the missile velocity is several times larger than that of the target)

--development of a rapid method for determining whether the control angles are increasing or decreasing

In the case of the method based upon power series expansions, the following things should be considered:

-- imposition of bounds upon the control variables ($\dot{\gamma}_A$ and $\dot{\gamma}_E$)

-- imposition of bounds upon the rates of change of the control variables

--study of the number of terms required in the power series expansions and derivation of all necessary terms

Long range, less well-defined goals, for one or both guidance methods include the following:

--maximize the probability of a kill assuming that errors with known distributions exist in the parameters defining the time-history of the position of the target

--remove the penalty term $\int_0^T |\dot{\vec{\Gamma}}|^2 dt$ from the performance index and bound $\dot{\vec{\Gamma}}$ in a more realistic manner (this might include introduction of a condition such as $V_M(t) = V_{NOMINAL}^{(t)} - \int_0^T \dot{\vec{\Gamma}}^T A(t) \dot{\vec{\Gamma}} dt$ which takes into account the degradation of velocity due to large changes in $\dot{\vec{\Gamma}}$)

--introduce some of the six-degree-of-freedom dynamics or confine motion to, say, the pitch plane, and take the pitch angle of the missile into account

--consider the interface of the guidance systems with the control systems of air-to-air missiles

--ascertain relative importance of measurements used to predict the motions of the target and the missile

APPENDIX: EXPRESSION OF INTEGRALS IN TERMS OF ELLIPTIC INTEGRALS

The integrals I_1 and J_c will be expressed in terms of elliptic integrals for which there are convenient series expansions and tables.

Supposing that the magnitude of the velocity of the missile is several times that of the target at the final time, the condition $\vec{R}_F \cdot \dot{\vec{R}}_F = 0$ implies $\vec{R}_F \cdot \vec{V}_{MF} \doteq 0$, where $\vec{V}_M$ is the velocity of the missile. Therefore $\eta_F \doteq \pm \pi/2$ because η is the angle between $\vec{R}_F$ and $\vec{V}_M$.

The condition $\eta_F \doteq \pm \pi/2$ motivates a formulation of the problem based upon an expansion of η_F about $\pi/2$. However, such a formulation will not be carried out in this report.

Assume $0 \leq \eta \leq \pi$ or $-\pi \leq \eta \leq 0$; in other words, η is not allowed to vary more than about $\pm \pi/2$ radians from η_F. This restricts the variation in γ_A to about $\pm \pi/2$ radians.

The assumptions stated above are made in order that the following formulation will be simplified. However, the formulation can be generalized to cover more general cases.

As stated earlier in the report, it is necessary that $\cos\eta_F - \cos\eta \geq 0$. If $\eta \in [0, \pi]$, this implies $\eta_F \leq \eta$ because $\cos\eta$ is a decreasing function on the interval $[0, \pi]$. Therefore, for $\eta \in [0, \pi]$, η and γ_A are decreasing with time. If $\eta \in [-\pi, 0]$, then γ_A is increasing. A preliminary analysis can determine whether γ_A is increasing or decreasing. If γ_A is increasing, the z-axis can be reversed by π radians in order that γ_A will be decreasing. Therefore, without loss of generality, we can confine ourselves to the case of $\eta \in [0, \pi]$, $\eta_F \doteq \pi/2$, and η decreasing with time.

Recall that

$$I_1 = \int_{\eta_0}^{\eta_F} \frac{d\eta}{\sqrt{\cos\eta_F - \cos\eta}}$$

Let $s = \cos\eta$. Then $ds = -\sin\eta\, d\eta$ where $\sin\eta = \sqrt{1-\cos^2\eta} = \sqrt{1-s^2}$ because $\sin\eta \geq 0$ for $\eta \in [0, \pi]$. Therefore,

$$I_1 = -\int_{\cos\eta_0}^{\cos\eta_F} \frac{ds}{\sqrt{1-s^2}\sqrt{\cos\eta_F - s}}$$

$$I_1 = -\int_{x}^{\beta_2} \frac{ds}{\sqrt{(s-\beta_1)(s-\beta_2)(s-\beta_3)}}$$

where $x = \cos\eta_0$, $\beta_1 = 1$, $\beta_2 = \cos\eta_F$, $\beta_3 = -1$. According to some handbooks of mathematical functions, the latter integral is

$$I_1 = -\sqrt{2}\int_0^{\phi} \frac{d\theta}{\sqrt{1-\left(\frac{1+\cos\eta_F}{2}\right)\sin^2\theta}}$$

$$I_1 = -\sqrt{2}\, F\left(\phi, \sqrt{\frac{1+\cos\eta_F}{2}}\right)$$

where

$$\cos^2\phi = \frac{(1-\cos\eta_F)(1+\cos\eta_0)}{(1+\cos\eta_F)(1-\cos\eta_0)}$$

and $F(\phi, k)$ denotes an incomplete elliptic integral of the first kind. One can show that

$$\phi = \cos^{-1}\sqrt{\frac{(1-\cos\eta_F)(1+\cos\eta_0)}{(1+\cos\eta_F)(1-\cos\eta_0)}} = \cos^{-1}\left[\frac{(1+\cos\eta_0)\sin\eta_F}{(1+\cos\eta_F)\sin\eta_0}\right]$$

and $\phi \in [0, \pi/2]$.

It can be shown that

$$\frac{1+\cos\eta_F}{2} = \sin^2\left(\frac{\pi-\eta_F}{2}\right)$$

Therefore

$$I_1 = -\sqrt{2}\, F\left(\phi \setminus \frac{\pi-\eta_F}{2}\right)$$

where

$$F(\phi\setminus\alpha) = \int_0^{\phi} \frac{d\theta}{\sqrt{1-\sin^2\alpha\sin^2\theta}}$$

Often F is plotted as a function of α and ϕ. For a given value of ϕ, F increases as $\alpha=(\pi-\eta_F)/2$ increases $(0\le\alpha\le\pi/2)$. Therefore, F decreases as η_F increases assuming ϕ is constant. Also ϕ decreases as η_F increases, and F decreases as ϕ decreases. In summary, F decreases monotonically as η_F increases. Moreover, F increases as η_0 increases.

Now the integral

$$J = \int_{\pi/2}^{X} \frac{\cos\eta}{\sqrt{\cos\eta_F-\cos\eta}}\, d\eta$$

will be expressed in terms of elliptic integrals. Let $\theta=\eta-\pi/2$ and $X^*=X-\pi/2$. Then

$$J = -\int_0^{X^*} \frac{\sin\theta}{\sqrt{\cos\eta_F+\sin\theta}}\, d\theta$$

From tables of mathematical functions,

$$J = -\sqrt{2}\left[F\left(\sin^{-1}\sqrt{\frac{1-\sin X^*}{1+\cos\eta_F}} \setminus \frac{\pi-\eta_F}{2}\right) - 2E\left(\sin^{-1}\sqrt{\frac{1-\sin X^*}{1+\cos\eta_F}} \setminus \frac{\pi-\eta_F}{2}\right)\right.$$

where E is an elliptic integral of the second kind, and it is required that

$$-\sin^{-1}\cos\eta_F \le X^* \le \pi/2$$

$$\pi/2 - \sin^{-1}\cos\eta_F \le X \le \pi$$

Since

$$\begin{aligned}\sin^{-1}\cos\eta_F &= \sin^{-1}\left[\tfrac{\pi}{2} + (\eta_F - \tfrac{\pi}{2})\right] \\ &= \sin^{-1}\left[-\sin(\eta_F - \tfrac{\pi}{2})\right] \\ &= -(\eta_F - \tfrac{\pi}{2})\end{aligned}$$

we have the requirement

$$\eta_F \le x \le \pi$$

Since $x^* = -\cos x$,

$$J = G\left(\sin^{-1}\sqrt{\frac{1+\cos x}{1+\cos\eta_F}} \setminus \frac{\pi-\eta_F}{2}\right)$$

where G denotes $\sqrt{2}\,(2E-F)$. Therefore,

$$\int_{\pi/2}^{\eta_F} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\,d\eta = G\left(\pi/2 \setminus \frac{\pi-\eta_F}{2}\right)$$

and, assuming as we are that $\eta_0 \ge \eta_F$, we have

$$\int_{\pi/2}^{\eta_0} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\,d\eta = G\left(\sin^{-1}\sqrt{\frac{1+\cos\eta_0}{1+\cos\eta_F}} \setminus \frac{\pi-\eta_F}{2}\right)$$

Since

$$J_c = \int_{\eta_0}^{\eta_F} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\,d\eta ,$$

we have

$$J_c = \int_{\pi/2}^{\eta_F} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\,d\eta - \int_{\pi/2}^{\eta_0} \frac{\cos\eta}{\sqrt{\cos\eta_F - \cos\eta}}\,d\eta$$

$$J_c = G\left(\pi/2 \setminus \frac{\pi-\eta_F}{2}\right) - G\left(\sin^{-1}\sqrt{\frac{1+\cos\eta_0}{1+\cos\eta_F}} \setminus \frac{\pi-\eta_F}{2}\right)$$

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

SPREAD SPECTRUM ECCM PERFORMANCE.

CW JAMMING SIGNAL REDUCTION IN

SPREAD SPECTRUM RECEIVERS.

Prepared by:	Maurice J. Bouvier, Jr., Ph.D.
Academic Rank:	Associate Professor
Department and University:	Department of Electrical Engineering Louisiana State University
Assignment:	Eglin AFB AFATL Guided Weapons Air to Surface Midcourse Guidance
USAF Research Colleague:	Major C. D. DePriest
Date:	August 5, 1977
Contract No.:	F44620-75-c-0031

SPREAD SPECTRUM ECCM PERFORMANCE

CW JAMMING SIGNAL REDUCTION IN SPREAD SPECTRUM RECEIVERS

by

Maurice J. Bouvier, Jr.

ABSTRACT

The Class M Tactical Global Positioning System Guidance (TGPSG) receiver being developed by AFATL for improved midcourse guidance of air-launched tactical weapons exhibits a relatively high degree of jamming resistance due to the processing gain inherent in spread spectrum systems. This processing gain is examined in order to identify system parameters which may be modified for improved jamming resistance.

For wideband jamming increased jamming resistance is accomplished by further reduction of the post-correlation bandwidth. The most significant increase could come through the removal of navigation data from the satellite signal. For CW jamming, which is both more effective and more easily generated, jamming resistance can also be increased through jamming signal tracking and cancellation. One such technique is presented and evaluated. Characteristics of the jamming and spread spectrum signals are such that the jamming signal can essentially be completely eliminated in theory. Examination of the results from a more conservative viewpoint still indicates a significant performance improvement.

INTRODUCTION

The objective of the Tactical Global Positioning System Guidance (TGPSG) program at AFATL is improved midcourse guidance of air-launched tactical weapons through the use of the 24 satellite NAVSTAR Global Positioning System (GPS). The GPS, which is to be operational in mid-1980, will provide highly accurate, continuous position and velocity information on a worldwide, all-weather basis. The TGPSG philosophy is to employ a highly sophisticated GPS receiver and processor (designated the X Set) in the launch aircraft and a relatively unsophisticated GPS receiver (designated Class M) and low-cost inertial guidance system (LCIGS) in the weapon. The X Set assists the weapon's GPS system in satellite selection and tracking prior to launch; after launch, the Class M receiver periodically updates the inertial system so that the relatively high drift rates in the LCIGS are offset. If during flight a jamming signal renders the Class M receiver inoperative, the LCIGS, with latest update but high drift, performs the terminal guidance function. The closer to the target effective jamming occurs, the more accurate the terminal guidance, which translates to a desire for improved jamming resistance by the Class M receiver. The identification and analytic evaluation of improved anti-jamming techniques which can be incorporated into the Class M receiver system are the primary objectives of this effort.

Class M receivers presently under development exhibit a relatively high degree of jamming resistance due to the processing gain inherent in spread spectrum systems. This processing gain is examined in the first part of this report. Although the results obtained are not original in themselves they are developed and presented here by such means as to unify and summarize those characteristics of particular importance in TGPSG system design. Results illustrate that jamming effectiveness is somewhat spectral shape sensitive with CW jamming slightly more effective than wideband jamming. Since CW jamming is also more easily generated improved ECCM performance against CW jamming is desirable. A technique which can potentially provide greatly increased CW jamming resistance is presented and analytically evaluated. The technique employs a phase-locked tracking loop in the early stages of the receiver to generate an estimate of the jamming signal and then uses the estimate to achieve jamming signal cancellation. Characteristics of the jamming signal and the satellite's spread spectrum signal are such that a very accurate estimate of the jamming signal can be ge[illegible] and hence highly effective cancellation achieved.

On the other hand wideband jamming resistance with no a priori knowledge of the signal characteristics must be accomplished primarily by the post-correlation narrowband processing. The narrowband processors incorporated in the present designs are known to perform near the theoretical limits so that little additional gain is likely to be achieved in this area. However further reduction of bandwidth by removal of the navigation date from one channel of the spread spectrum signal would provide a significant improvement in processing gain.

OBJECTIVES

The objectives of this effort are to identify and evaluate techniques for improved jamming resistanceof Class M TGPSG receivers primarily through electronic signal processing. One such technique which should prove effective against CW jamming has been identified and has bee n evaluated to a limited extent.

SPREAD SPECTRUM ECCM PERFORMANCE

The input to a direct sequence spread spectrum receiver typically consists of a spread spectrum signal $s_{ss}(t)$ of power S and a jamming signal j(t) of power J and can be expressed as

$$\begin{aligned} r(t) &= s_{ss}(t) + j(t) \\ &= \sqrt{2S}\, m(t)\, p(t) \cos \omega_c t \\ &\quad + \sqrt{2J}\, j_b(t) \cos \omega_c t \end{aligned} \tag{1}$$

where

m(t) = message waveform, bandlimited to W_m Hz, assumed unit power

p(t) = psuedo-random noise (PN) spectrum-spreading binary code, chip rate $1/T_p$

$j_b(t)$ = baseband representation of the jamming signal, assumed unit power

Demodulation is accomplished in the receiver by the operations indicated in Figure 1.

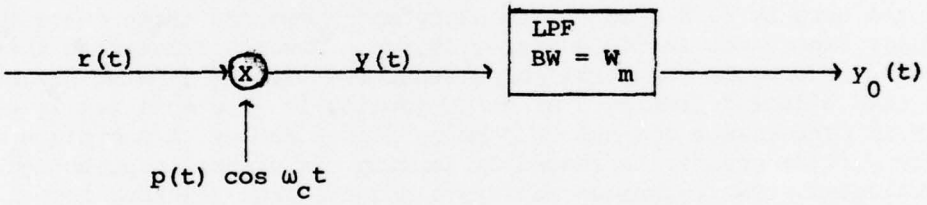


Figure 1. Demodulation of the Direct Sequence Spread Spectrum Signal

The multiplication produces

$$\begin{aligned} y(t) &= r(t) \left[p(t) \cos \omega_c t \right] \\ &= \sqrt{2S}\, m(t)\, p^2(t) \cos^2 \omega_c t \\ &\quad + \sqrt{2J}\, j_b(t)\, p(t) \cos^2 \omega_c t. \end{aligned}$$

Using the trig identity

$$\cos^2 X = \frac{1}{2}(1 + \cos 2X)$$

and the fact that

$$p(t) = \pm 1$$

so that

$$p^2(t) = 1$$

y(t) becomes

$$y(t) = \sqrt{\frac{S}{2}}\, m(t) + \sqrt{\frac{S}{2}}\, m(t) \cos 2\omega_c t + \sqrt{\frac{J}{2}}\, j_b(t)\, p(t) + \sqrt{\frac{J}{2}}\, j_b(t)\, p(t) \cos 2\omega_c t$$

Since the double frequency terms will not influence the output of the Low Pass Filter (LPF) they can be dropped from further consideration by defining

$$y_1(t) \triangleq y(t) - \{\text{Double Frequency Terms}\}$$

$$= \sqrt{\frac{S}{2}}\, m(t) + \sqrt{\frac{J}{2}}\, j_b(t)\, p(t) \qquad (2)$$

The message term is bandlimited to W_m so that it will pass through the LPF and produce an output signal with power

$$S_0 = \frac{S}{2}$$

The jamming signal term will produce an output jamming signal $j_0(t)$ with power J_0 which depend on the spectral characteristic of $j_b(t)\, p(t)$ in the bandpass of the LPF. The PN code waveform p(t) has power spectral density

$$P(f) = T_p \left(\frac{\sin \pi f T_p}{\pi f T_p}\right)^2 \qquad (3)$$

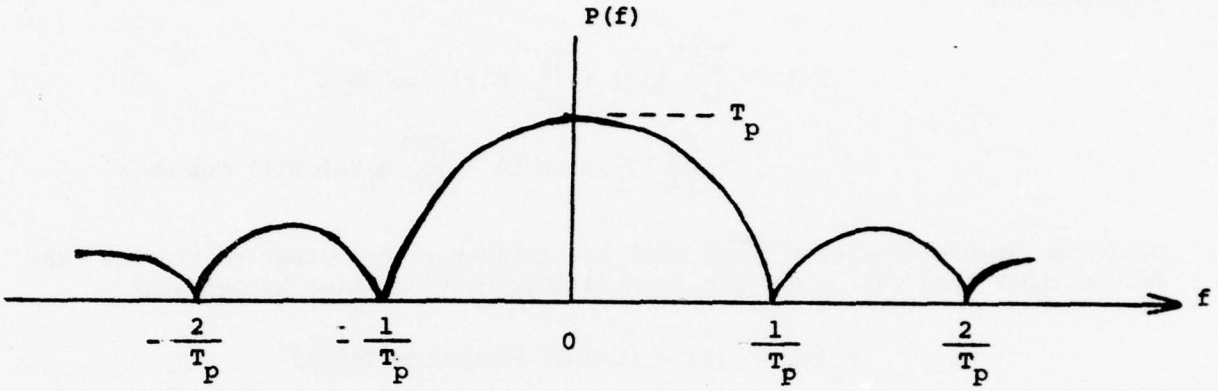
The jamming signal j(t) can be modelled as a wideband or narrowband PN signal or as a CW signal by selecting a power spectral density for $j_b(t)$ of

$$J_b(f) = T_j \left(\frac{\sin \pi f T_j}{\pi f T_j}\right)^2 \qquad (4)$$

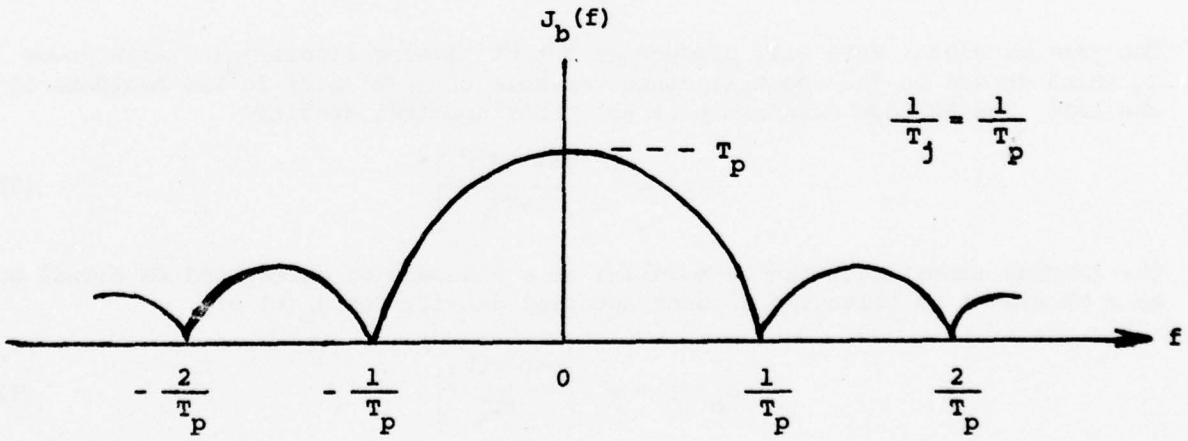
Wideband PN jamming corresponds to large chip rate $1/T_j$; narrowband PN jamming corresponds to small chip rate; CW jamming corresponds to zero chip rate in which

case $J_b(f)$ becomes a unit area impulse located at the origin. It is unlikely that $1/T_j$ would be greater than $1/T_p$ since that would spread the jamming signal power over a larger bandwidth than the spread spectrum signal bandwidth and would be an inefficient utilization of jammer power. Thus the chip width T_j can be expected to take on values in the range T_p to ∞ with $T_j = T_p$ corresponding to wideband jamming and $T_j = \infty$ corresponding to CW jamming. The PN code spectrum and three jamming signal spectra of interest are illustrated in Figure 2.

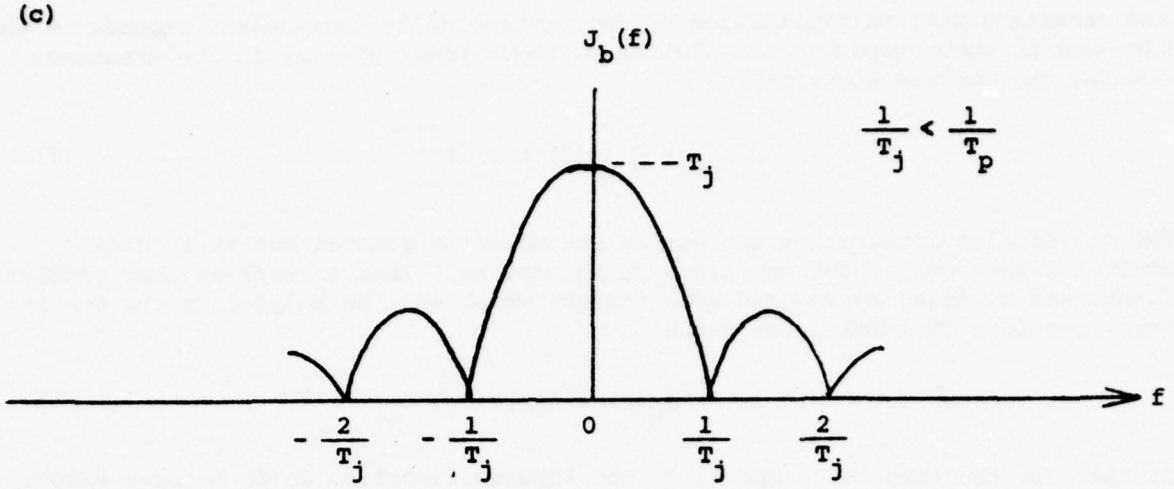
(a)



(b)



(c)



(d)

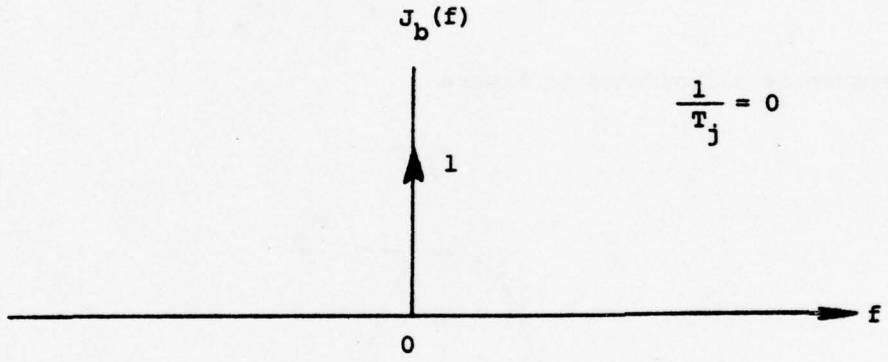


Figure 2. PN Code and Jamming Signal Spectra (a) PN Code (b) Wideband Jamming (c) Narrowband Jamming (d) CW Jamming

The objective then is to determine the spectrum of the jamming signal term in (2) so as to determine the amount of jamming signal power which passes through the LPF. Defining

$$j_1(t) \triangleq \sqrt{\frac{J}{2}}\, J_b(t)\, p(t) \tag{5}$$

and recalling that multiplication of two statistically independent signals in the time domain corresponds to convolution of their power spectra in the frequency domain, the desired spectrum is

$$J_1(f) = \frac{J}{2}\left[p(f) * J_b(f)\right] \tag{6}$$

The convolution integral is not easily evaluated in general but it is easily evaluated when one of the functions is an impulse. Thus to work an easy problem first, and by doing so to gain some insight which will be helpful in the general case, consider CW jamming for which

$$J_b(f) = \delta(f)$$

so that, by the integral property of the impulse function, $J_1(f)$ becomes simply

$$J_1(f) = \frac{J}{2}\ p(f)$$

$$= \frac{JT_p}{2}\left(\frac{\sin \pi f T_p}{\pi f T_p}\right)^2 \tag{7}$$

This spectrum is illustrated in Figure 3.

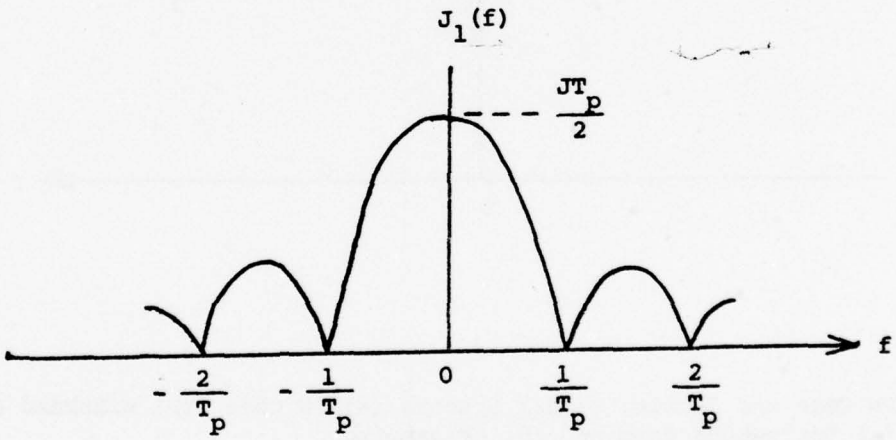


Figure 3. Jamming Signal Spectrum at Input to LPF. CW Jamming.

The output jamming signal $j_0(t)$ has a spectrum which is the same as $J_1(f)$ over the bandpass of the LPF and zero elsewhere, i.e.,

$$J_0(f) = \begin{cases} J_1(f) , & |f| \leq W_m \\ 0 , & |f| > W_m \end{cases}$$

In spread spectrum systems the chip rate $1/T_p$ is many orders of magnitude larger than the message bandwidth W_m so that the output jamming signal spectrum is as shown in Figure 4 but where the relative values of $1/T_p$ and W_m are not accurately portrayed (may typically be 10^6 to 1).

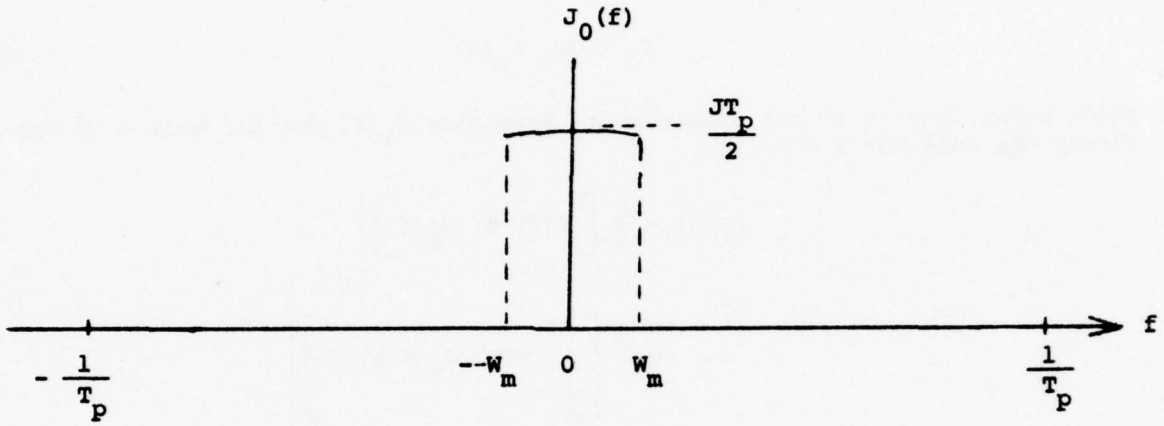


Figure 4. Output Jamming Signal Spectrum. CW Jamming.

With the typically large ratios of $1/T_p$ to W_m the output jamming signal spectrum can be assumed to be constant over the $-W_m$ to W_m range and the output jamming power well approximated by

$$J_0 = J\ T_p\ W_m \tag{8}$$

and the output jamming power to message power ratio by

$$\frac{J_0}{S_0} = \frac{J\ T_p W_m}{S/2}$$

$$= \frac{J}{S}\ 2T_p W_m$$

$$= 2T_p W_m \left[\frac{J}{S}\right)_{\text{input}}\Bigg]$$

$$\left.\frac{J}{S}\right)_{\text{output}} = \frac{2\ W_m}{1/T_p}\left[\frac{J}{S}\right)_{\text{input}}\Bigg] \tag{9}$$

For the general case the convolution of (6)

$$J_1(f) = \frac{J}{2}\left[P(f) \ast J_b(f)\right]$$

must be considered for the general $J_b(f)$ spectrum. But convolution of two wideband spectra will produce another wideband spectrum which is quite flat in the vicinity of the origin. Due to this flatness at the origin of $J_1(f)$, the small bandwidth of the LPF and the large bandwidth of $J_1(f)$, the output jamming power can again be well approximated by

$$J_0 = 2W_m J_1(0) \tag{10}$$

which means that it is not necessary to determine $J_1(f)$ for all values of frequency but only for $f = 0$, i.e.,

$$\begin{aligned} J_1(0) &= \frac{J}{2}\left[P(f) \ast J_b(f)\right]\Big|_{f=0} \\ &= \frac{J}{2}\left[\int_{-\infty}^{\infty} P(\mu)\, J_b(f-\mu)\, du\right]\Bigg|_{f=0} \\ &= \frac{J}{2}\int_{-\infty}^{\infty} P(\mu)\, J_b(-\mu)du \\ &= \frac{J}{2}\int_{-\infty}^{\infty} P(\mu)\, J_b(\mu)\, du. \end{aligned}$$

Although evaluation of this integral for spectra of interest would provide the desired spectral value, it is easier and somewhat more enlightening to work with the corresponding autocorrelation functions, $R_p(\tau)$ and $R_b(\tau)$, the inverse Fourier transforms of $P(f)$ and $J_b(f)$ respectively. For the spectra of interest, given by (3) and (4), the autocorrelation functions are

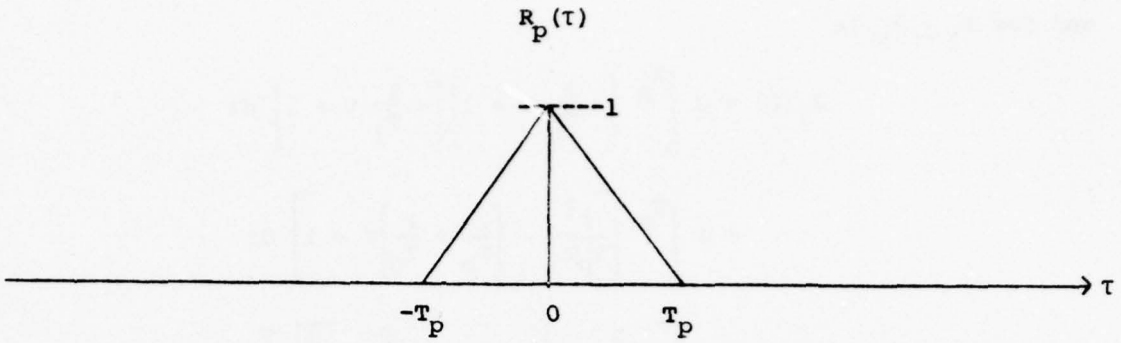
$$R_p(\tau) = \begin{cases} -\frac{1}{T_p}|\tau| + 1, & |\tau| \le T_p \\ 0, & |\tau| > T_p \end{cases}$$

and

$$R_b(\tau) = \begin{cases} -\frac{1}{T_j}|\tau| + 1, & |\tau| \le T_j \\ 0, & |\tau| > T_j \end{cases}$$

which are illustrated in Figure 5.

(a)



(b)

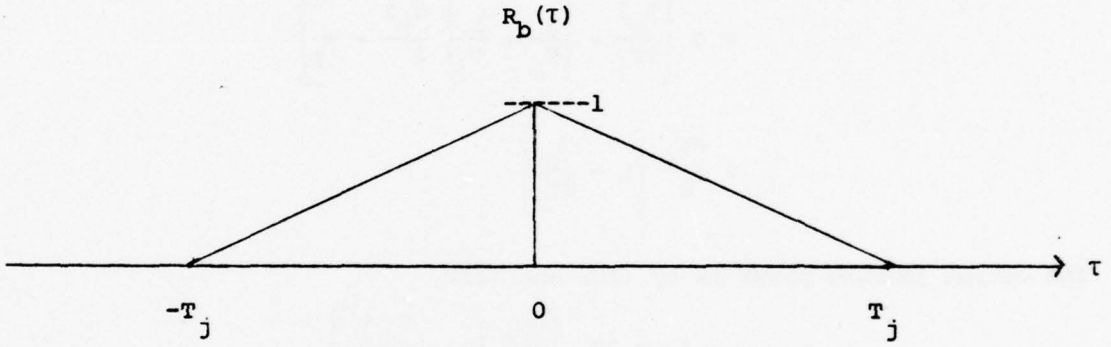


Figure 5. Autocorrelation Functions Corresponding to (a) P(f), (b) $J_b(f)$.

$J_1(f)$ is then proportional to the Fourier transform of the product of $R_p(\tau)$ and $R_b(\tau)$,

$$J_1(f) = \frac{J}{2} \int_{-\infty}^{\infty} R_p(\tau)\ R_b(\tau)\ e^{-j\omega\tau} d\tau$$

But the value of interest is $J_1(0)$

$$J_1(0) = \frac{J}{2} \int_{-\infty}^{\infty} R_p(\tau)\ R_b(\tau)\ d\tau$$

which due to the symmetry of $R_p(\tau)$ and $R_b(\tau)$ can be evaluated by

$$J_1(0) = J \int_0^\infty R_p(\tau)\, R_b(\tau)\, d\tau$$

and for $T_j \geq T_p'$ is

$$J_1(0) = J \int_0^{T_p} \left(-\frac{1}{T_p}\tau + 1\right)\left(-\frac{1}{T_j}\tau + 1\right) d\tau$$

$$= J \int_0^{T_p} \left[\frac{\tau^2}{T_p T_j} - \left(\frac{1}{T_p} + \frac{1}{T_j}\right)\tau + 1\right] d\tau$$

$$= J \left[\frac{\tau^3}{3T_p T_j} - \left(\frac{1}{T_p} + \frac{1}{T_j}\right)\frac{\tau^2}{2} + \tau\right]\Bigg|_0^{T_p}$$

$$= J \left[\frac{T_p^2}{3T_j} - \left(\frac{1}{T_p} + \frac{1}{T_j}\right)\frac{T_p^2}{2} + T_p\right]$$

$$= \frac{JT_p}{2}\left(1 - \frac{T_p}{3T_j}\right) \tag{11}$$

The output jamming power is by (10) and (11)

$$J_0 = 2W_m \left[\frac{JT_p}{2}\left(1 - \frac{T_p}{3T_j}\right)\right]$$

$$= JT_p\, W_m \left(1 - \frac{T_p}{3T_j}\right) \tag{12}$$

and the output jamming power to signal power ratio is

$$\frac{J_0}{S_0} = \frac{J\, T_p W_m \left(1 - \frac{T_p}{3T_j}\right)}{S/2}$$

or

$$\boxed{\left.\frac{J}{S}\right)_{\text{output}} = 2\,\frac{W_m}{1/T_p}\left(1 - \frac{T_p}{3T_j}\right)\left[\left.\frac{J}{S}\right)_{\text{input}}\right]} \tag{13}$$

For CW jamming, $T_j = \infty$ and

$$\left(\frac{J}{S}\right)_{\text{output}} = 2\,\frac{W_m}{1/T_p}\left[\left(\frac{J}{S}\right)_{\text{input}}\right]$$

CW Jamming

as was found for this special case earlier. For wideband jamming $T_j = T_p$ and

$$\left(\frac{J}{S}\right)_{\text{output}} = \frac{4}{3}\,\frac{W_m}{1/T_p}\left[\left(\frac{J}{S}\right)_{\text{input}}\right] \tag{14}$$

Wideband Jamming

Comparing these two results indicates that CW jamming is 50% more effective than wideband jamming, that is, for the same jamming power

$$\left(\frac{J}{S}\right)_{\substack{\text{output}\\ \text{CW}}} = 1.5\left[\left(\frac{J}{S}\right)_{\substack{\text{output}\\ \text{WB}}}\right] \tag{15}$$

Note from (14) that the JSR is reduced from input to output be a factor of

$$\frac{4}{3}\,\frac{W_m}{1/T_p}$$

or a "processing gain" of

$$G_p = \frac{3}{4}\,\frac{1/T_p}{W_m} \cong \frac{1/T_p}{W_m} \tag{16}$$

is achieved. If the message m(t) is a digital waveform with pulse width T_m then

$$W_m \cong \frac{1}{T_m}$$

and the processing gain is

$$G_p \cong \frac{1/T_p}{1/T_m}$$

$$G_p = \frac{\text{PN Chip Rate}}{\text{Message Chip Rate}} \qquad (17)$$

An idealized case often considered is with both the PN code p(t) and the baseband jamming signal $j_b(t)$ having rectangular spectra as indicated in Figure 6, where $W_j \leq W_p$. (Note that $W_j = 0$ corresponds to CW jamming).

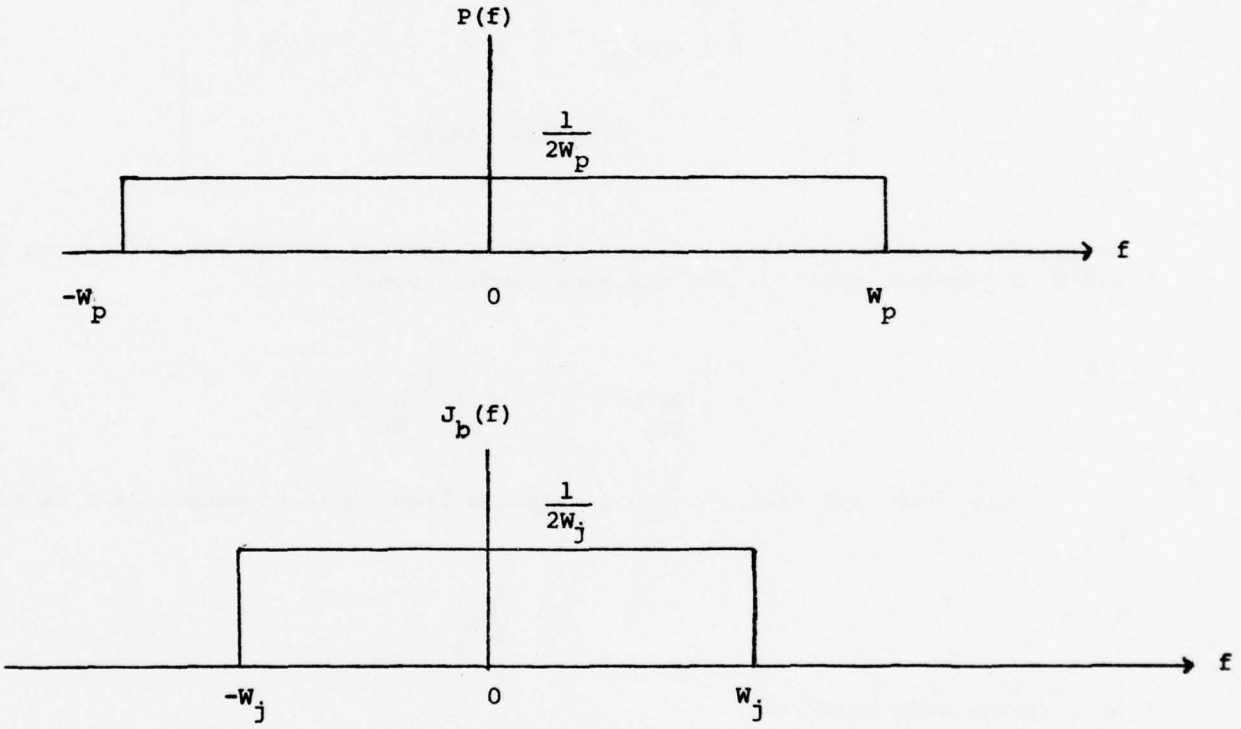


Figure 6. Rectangular Spectra for p(t) and $j_b(t)$.

$J_1(f)$ is readily determined in this case by direct convolution of P(f) and $J_b(f)$ with the result as shown in Figure 7.

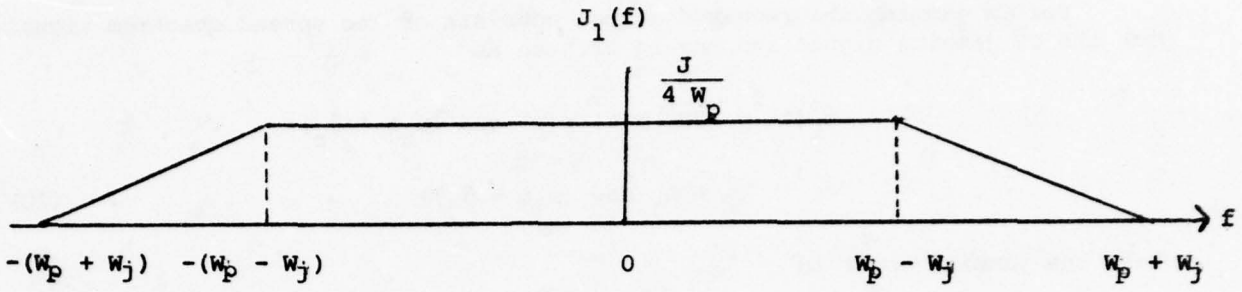


Figure 7. Jamming Signal Spectrum at Input to LPF. Rectangular p(t) and $j_b(t)$ Spectra.

The jamming power in the LPF bandpass is then

$$J_0 = \frac{W_m J}{2 W_p} \tag{18}$$

so that the output JSR is

$$\frac{J_0}{S_0} = \frac{W_m}{W_p} \frac{J}{S}$$

and the processing gain is

$$\boxed{Gp = \frac{W_p}{W_m}} \tag{19}$$

which is the same as for the general case given by (17).

It is interesting to note that if the PN code and jamming signal spectra are rectangular the output jamming power given by (18) is independent of the jamming signal's bandwidth. Thus for this case <u>wideband jamming is just as effective as CW jamming</u>.

It is also important to note that what has been considered to be a jamming signal may in fact be another spread spectrum signal operating on the same frequency but with a different PN code. The second spread spectrum signal's interference effects are reduced by the processing gain of the receiver so that spread spectrum communication systems tend to be interference free if all transmitters are of comparable power.

CW JAMMING SIGNAL REDUCTION IN SPREAD SPECTRUM RECEIVERS

For CW jamming the received signal consists of the spread spectrum signal plus the CW jamming signal and may be written as

$$r_1(t) = \sqrt{2S}\, m(t)\, p(t) \cos(\omega_c t + \theta_c) + A_j \cos(\omega_j t + \theta_j) \tag{20}$$

where the jamming power is

$$Pj = \frac{A_j^2}{2} .$$

For the case of high JSR the received signal consists of a large power jamming signal concentrated at one frequency and a small power spread spectrum signal spread over a large frequency range so that a narrowband tracking circuit can produce a very accurate estimate of the jamming signal. This estimate of the jamming signal can then be used to cancel the jamming signal prior to correlation processing in the receiver. One possible estimation technique is illustrated in Figure 8. The net result is an input to the receiver of

$$r_2(t) = s_{ss}(t) + A_j \cos(\omega_j t + \theta_j) - \hat{A}_j \cos(\omega_j t + \hat{\theta}_j) \tag{21}$$

so that the effective jamming signal is

$$j_2(t) = A_j \cos(\omega_j t + \theta_j) - \hat{A}_j \cos(\omega_j t + \hat{\theta}_j) \tag{22}$$

Using the trig identity for the sum of two sinusoids of the same frequency

$$A \cos(\omega t + \alpha) + B \cos(\omega t + \beta)$$

$$= C \cos(\omega t + \gamma)$$

where

$$C = \sqrt{A^2 + B^2 + 2AB \cos(\alpha - \beta)}$$

and

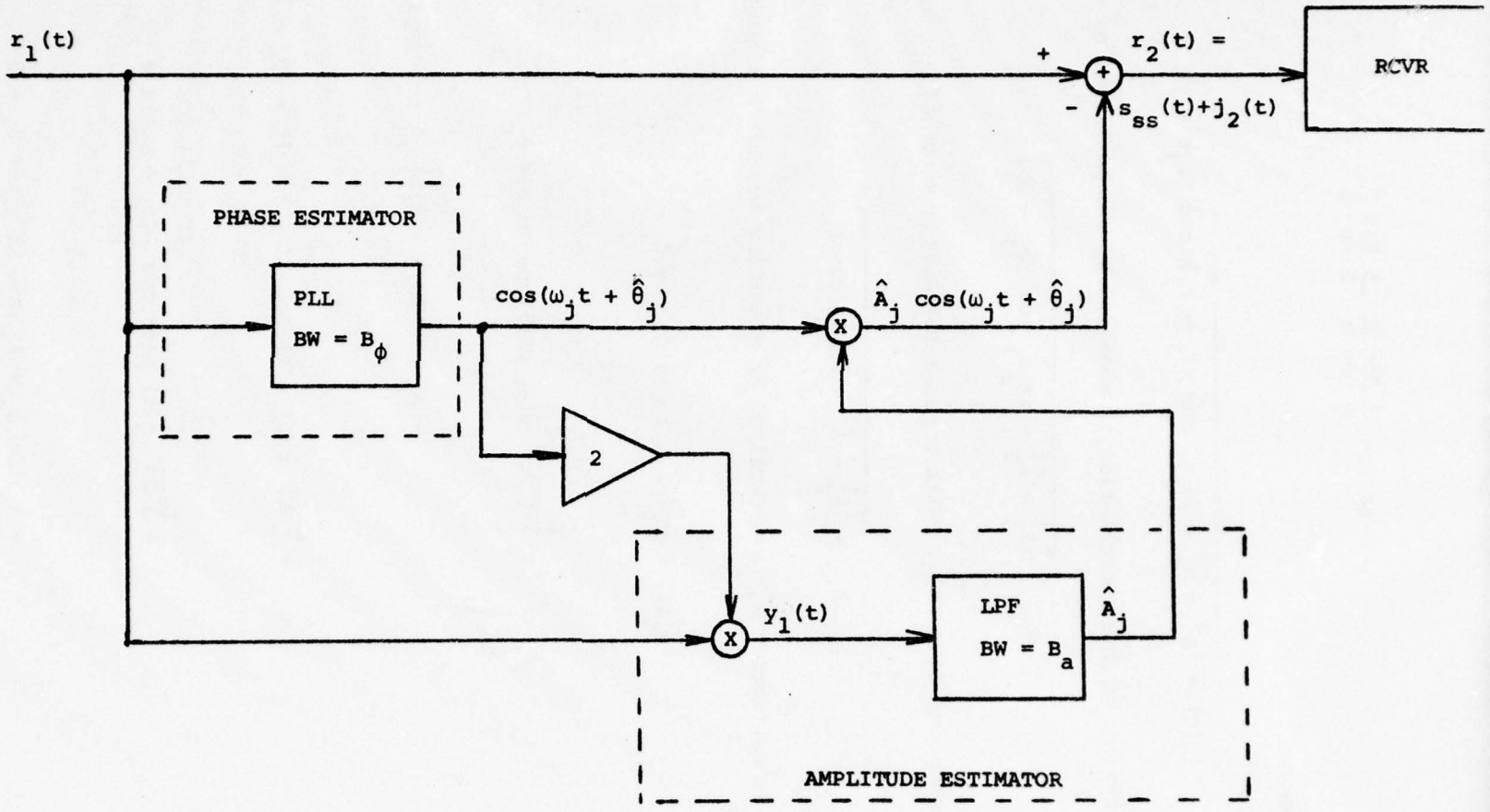


Figure 8. CW Jamming Signal Reduction Block Diagram

$$\gamma = \arctan \frac{A \sin \alpha + B \sin \beta}{A \cos \alpha + B \cos \beta}$$

$j_2(t)$ becomes

$$j_2(t) = \sqrt{A_j^2 + \hat{A}_j^2 - 2A_j\hat{A}_j \cos (\theta_j - \hat{\theta}_j)} \cos (\omega_j t + \gamma) \tag{23}$$

The effectiveness of the cancellation is measured by the amplitude of $j_2(t)$

$$A_{jc} \triangleq \sqrt{A_j^2 + \hat{A}_j^2 - 2A_j\hat{A}_j \cos (\theta_j - \hat{\theta}_j)}$$

or its mean-squared value $\overline{A_{jc}^2}$. With a phase estimation error of ϕ_e, A_{jc} is

$$A_{jc} = \sqrt{A_j^2 + \hat{A}_j^2 - 2A_j\hat{A}_j \cos \phi_e} \tag{24}$$

The amplitude estimation $\hat{A}_j$ is determined by examining the LPF's response to $y_1(t)$ where

$$y_1(t) = r_1(t) \left[2 \cos (\omega_j t + \hat{\theta}_j\right]$$

$$= \left[\sqrt{2S}\, m(t)\, p(t) \cos (\omega_c t + \theta_c) + A_j \cos (\omega_j t + \theta_j)\right] \left[2 \cos (\omega_j t + \hat{\theta}_j)\right]$$

$$= \sqrt{2S}\, m(t)\, p(t) \cos \left[(\omega_c - \omega_j)t + (\theta_c - \hat{\theta}_j)\right]$$

$$+ \sqrt{2S}\, m(t)\, p(t) \cos \left[(\omega_c + \omega_j)\, t + (\theta_c + \hat{\theta}_j)\right]$$

$$+ A_j \cos \phi_e + A_j \cos (2\omega_j t + \theta_j + \hat{\theta}_j).$$

The second and fourth terms will not pass the LPF so define

$$y_2(t) \triangleq A_j \cos \phi_e$$

$$+ \sqrt{2S}\, m(t)\, p(t) \cos \left[(\omega_c - \omega_j)t + (\theta_c - \hat{\theta}_j) \right] \quad (25)$$

With ϕ_e small the first term represents a good estimate of A_j whereas the second term represents the "noise" in the amplitude estimate at the input to the LPF. To determine the corresponding noise at the LPF output define

$$n_2(t) \triangleq \sqrt{2S}\, m(t)\, p(t) \cos \left[(\omega_c - \omega_j)\, t + (\theta_c - \hat{\theta}_j) \right]$$

whose power spectrum is

$$N_2(f) = \frac{S}{2} \left[M(f) \circledast P(f) \right] \circledast \left[\delta(f - f_c + f_j) + \delta(f + f_c - f_j) \right] \quad (26)$$

With m(t) very narrowband compared to p(t),

$$M(f) \circledast P(f) \cong P(f)$$

and for $f_j \cong f_c$, which is a worst case condition, $N_2(f)$ becomes

$$N_2(f) = S\, P(f) \circledast \delta(f)$$

$$= S\, P(f)$$

Since the LPF bandwidth B_a is very small compared to the bandwidth of p(t), the noise at the output of the LPF, $n_0(t)$, has power

$$N_0 = 2\, B_a\, N_2(0)$$

$$= 2\, B_a S\, P(0)$$

$$= 2\, B_a S\, T_p$$

or since

$$W_{ss} \cong \frac{2}{T_p} ,$$

$$N_0 = \frac{4 B_a S}{W_{ss}} \tag{27}$$

Returning to A_{jc} of (24)

$$A_{jc} = \sqrt{A_j^2 + \hat{A}_j^2 - 2A_j \hat{A}_j \cos \phi_e}$$

$$= \sqrt{A_j^2 + (A_j \cos \phi_e + n_0)^2 - 2A_j (A_j \cos \phi_e + n_0) \cos \phi_e}$$

$$= \sqrt{A_j^2 + A_j^2 \cos^2 \phi_e + 2n_0 A_j \cos \phi_e + n_0^2 - 2A_j^2 \cos^2 \phi_e - 2n_0 A_j \cos \phi_e}$$

$$= \sqrt{A_j^2 - A_j^2 \cos^2 \phi_e + n_0^2}$$

$$= \sqrt{A_j^2 - \frac{A_j^2}{2} - \frac{A_j^2}{2} \cos 2\phi_e + n_0^2}$$

$$= \sqrt{\frac{A_j^2}{2} - \frac{A_j^2}{2} \cos 2\phi_e + n_0^2}$$

Expanding the cosine term in a power series

$$\cos 2\phi_e = 1 - \frac{(2\phi_e)^2}{2!} + \frac{(2\phi_e)^4}{4!} - \ldots$$

and assuming that

$$2\phi_e << 1 ,$$

as will be shown later, so that $\cos 2\phi_e$ is well represented by the first two terms,

$$A_{jc} = \sqrt{\frac{A_j^2}{2} - \frac{A_j^2}{2} + \frac{A_j^2}{2}\frac{(2\phi_e)^2}{2} + n_0^2}$$

$$= \sqrt{A_j^2 \phi_e^2 + n_0^2}$$

and its mean-squared value is

$$\overline{A_{jc}^2} = A_j^2 \overline{\phi_e^2} + \overline{n_0^2}$$

$$= A_j^2 \overline{\phi_e^2} + N_0$$

$$\boxed{\overline{A_{jc}^2} = 2J \overline{\phi_e^2} + N_0} \qquad (28)$$

where N_0 is given by (27).

Viterbi [1] shows that for a PLL with input consisting of $\sqrt{2}\,A \cos(\omega_0 t + \theta)$ plus white noise of double-sided spectral density n/2 and loop bandwidth B_L the variance of the phase error is

$$\sigma_{\phi_e}^2 = \frac{n B_L}{A^2}$$

For our case the input is $A_j \cos(\omega_j t + \theta_j)$ plus the wideband spread spectrum signal of power S and bandwidth W_{ss} which to the narrowband PLL looks like white noise of spectral density $\frac{S}{2W_{ss}}$ so that the error variance is

$$\sigma_{\phi_e}^2 = \frac{2S B_\phi}{A_j^2 W_{ss}}$$

$$= \frac{S B_\phi}{J W_{ss}}$$

If the phase estimate is unbiased

$$\overline{\phi_e^2} = \sigma_{\phi_e}^2$$

$$= \frac{S\ B_\phi}{J\ W_{ss}} \tag{29}$$

Note that for a JSR as small as 0 db

$$\overline{\phi_e^2} = \frac{B_\phi}{W_{ss}}$$

which may be as small as 10^{-6}, so that the earlier assumption that $2\ \phi_e << 1$ is well justified. With (27) and (29) in (28)

$$\overline{A_{jc}^2} = 2J \left[\frac{S\ B_\phi}{J\ W_{ss}} \right] + \frac{4\ B_a S}{W_{ss}}$$

$$= 2\ \frac{B_\phi + 2\ B_a}{W_{ss}}\ S$$

which corresponds to jamming signal power after cancellation of

$$P_{jc} = \frac{\overline{A_{jc}^2}}{2}$$

$$\boxed{P_{jc} = \frac{B_\phi + 2\ B_a}{W_{ss}}\ S} \tag{30}$$

and a JSR of

$$\boxed{\frac{P_{jc}}{S} = \frac{B_\phi + 2\ B_a}{W_{ss}}} \tag{31}$$

Thus the CW jamming signal reduction circuit exhibits performance analogous to the processing gain of a spread spectrum receiver. This is not unexpected though since the estimator is operating on a narrowband signal (the CW jamming signal) in wideband noise (the spread spectrum signal). It is interesting to note that the power in the jamming signal after cancellation is independent of the received jamming signal's power. This is due to the fact that as the incoming jamming signal increases in strength the estimator's performance improves thereby providing more effective cancellation. Of course for very small jamming signals the estimator's performance degrades and it should therefore be disabled.

To get a feel for the effectiveness of the jamming cancellation assume

(1) $B_\phi = B_a = 100$ Hz

(2) $W_{ss} = 20$ MHz

(3) $S = -160$ dbw $= 10^{-16}$ w

so that the effective jamming power is

$$P_{jc} = 1.5 \times 10^{-21} \text{ w}$$

$$\cong -208 \text{ db}$$

and the effective JSR is

$$\frac{P_{jc}}{S} = 1.5 \times 10^{-5}$$

$$\cong -48 \text{ db}$$

and both are independent of the received jamming signal power.

It should be pointed out that the highly effective cancellation indicated here cannot be achieved in practice due to such factors as implementation errors, system time delays and non-linearities, "sloppiness" of the CW jamming signal, RF leakage, etc. A more conservative figure may be obtained by writing (28) in terms of power

$$P_{jc} = \frac{\overline{A_{jc}^2}}{2}$$

$$= J \overline{\phi_e^2} + \frac{\overline{n_0^2}}{2}$$

and assuming

(1) rms value of $\phi_e = 0.1$ rad

(2) rms value of $n_0 = 0.1 A_j$

so that

$$P_{jc} = J\ (0.01) + \frac{0.01\ A_j^2}{2}$$

$$= J\ (0.01) + (0.01)\ J$$

$$= 0.02\ J$$

which is a 17 db reduction in jamming power--nowhere near the theoretical value but still a significant amount.

CONCLUSIONS AND RECOMMENDATIONS

The GPS system's spread spectrum feature provides a considerable processing gain against both CW and wideband jamming. Increased jamming resistance through this feature requires a further decrease in post-correlation bandwidth (or increase in PN code chip rate which is unlikely). Although some improvement is possible through the techniques discussed in [2] the greatest bandwidth reduction and hence the most dramatic performance improvement would come from removal of the navigation data from the GPS signal. Simulation studies in [2] indicate a 6 to 10 db gain by data removal so that this avenue should be pursued. Of course there are other means of reducing JSR such as increasing transmitter power or incorporating an adaptive antenna pattern capability at the receiver, the latter of which should be pursued.

Examination of the spread spectrum processing gain indicates that CW jamming is somewhat more effective than wideband jamming. Since CW jamming is also more easily generated, the potential advantage afforded by increased CW jamming resistance is obvious. A technique which can potentially provide greatly increased CW jamming resistance is presented and analyzed. Characteristics of the jamming and spread spectrum signals are such that the jamming signal can essentially be completely eliminated in theory. Examination of the results from a more conservative point of view still indicates a significant performance advantage. Further evaluation of this technique is recommended with particular attention directed towards implementation techniques and errors and to its effectiveness against "sloppy" jamming signals, that is, where the amplitude, frequency and phase are randomly time-varying functions.

REFERENCES

1. Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill, 1966, p. 35.

2. Brammer, R. F. and J. G. Himes, "Evaluation of jamming resistant GPS receivers with application to tactical missile guidance," The Analytical Sciences Corp., Reading, MA, Report No. TR-625-3, 15 August 1976.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY

PARTICIPANT'S FINAL REPORT

AN ANALYSIS OF AERODYNAMIC FORCE AND MOMENT MODELS FOR ASYMMETRIC MUNITIONS

Prepared by:	Robert W. Courter, PhD
Academic Rank:	Associate Professor
Department and University:	Department of Mechanical Engineering Louisiana State University
Assignment:	
(Air Force Base) (Laboratory) (Division) (Branch)	Eglin AFB Armament Laboratory Guns, Rockets and Explosives Ballistics
USAF Research Colleague:	Kenneth K. Cobb
Date:	August 19, 1977
Contract No.:	F44620-75-C-0031

AN ANALYSIS OF AERODYNAMIC FORCE AND MOMENT MODELS FOR ASYMMETRIC MUNITIONS

by

Robert W. Courter

ABSTRACT

A study has been made of the dynamic and aerodynamic models currently being used in aerodynamic coefficient extraction programs at the Aeroballistic Research Facility at Eglin AFB, Florida. The objectives of the study were to review the extraction programs and aerodynamic models to determine what alterations could be made to improve their accuracy in predicting coefficients for unsymmetrical munitions.

Two analysis programs are available, one structured to deal with projectiles having multiple planes of symmetry and the other designed to handle configurations with only one plane of symmetry. The present study was involved primarily with the latter program. It was found that the formulation of the dynamic equations did not account generally for c.g. axis offset and, hence, limited the applicability of the program to configurations with very small offsets. A set of transformation equations was derived to alleviate this problem.

Studies of the aerodynamic models contained within the program indicate that they are restricted to certain ranges of flight angle and to certain types of configurations. Comparisons of these models with wind tunnel data suggest possible methods of modification which would yield better results.

It is recommended that a thorough study of aerodynamic modeling with appropriate ballistic test support be initiated to improve the extraction models so that existing configurations can be tested with confidence and configuration limits for ballistic testing can be assessed.

ACKNOWLEDGEMENTS

The author appreciates the opportunity to participate in this exceptional summer research program, and he commends the Air Force Systems Command for its wisdom in establishing this valuable communication link between the universities and the Air Force. Thanks are also given to the ASEE and to Auburn University for handling the administration of the program. Mr. Fred O'Brien deserves special praise in this regard for performing his administrative function so well that the author could devote all of his time to research.

The Ballistics Branch of the Guns, Rockets and Explosives Division has proven to be an ideal group in which to do productive work. Division Chief Col Gerald D'Arcy, Branch Chief Col Donald Chomicz, and Technical Advisor Dr. Donald Daniel have all expended every effort to make the author's stay both comfortable and technically rewarding.

The summer has been spent in the company of several engineers who have tirelessly answered questions, provided data and served as a sounding board for ideas as they have evolved. Foremost among these is Mr. Kenneth Cobb who, in his capacity as Research Colleague, suggested and guided the research with aplomb. Other branch members who deserve thanks are Lt Demo Galanos, Mrs. Judith Sullivan Kleist, Mr. Kenneth West and Mr. Jerry Winchenbach. The author is also grateful to Miss Emma Walmer who typed the manuscript.

NOMENCLATURE

A	Reference Area
B	Asymmetries Euler angle Matrix
b_{ij}	Matrix elements of B
C	Inertia transformation matrix
c_{ij}	Matrix elements of C
C_{l_β}	$\frac{\partial (L/\bar{q}_b A \ell)}{\partial \beta}$
$C_{m_{\dot{\alpha}}}$	$\frac{\partial (M/\bar{q}_b A \ell)}{\partial (\dot{\alpha} \ell/2V)}$
C_{n_β}	$\frac{\partial (N/\bar{q}_b A \ell)}{\partial \beta}$
C_{X_α}	$\frac{\partial (F_x/\bar{q}_b A)}{\partial \alpha}$
$C_{X_{\dot{\alpha}}}$	$\frac{\partial (F_x/\bar{q}_b A)}{\partial (\dot{\alpha} \ell/2V)}$
C_{X_u}	$\frac{\partial (F_x/\bar{q}_b A)}{\partial (u/V)}$
C_{y_β}	$\frac{\partial (F_y/\bar{q}_b A)}{\partial \beta}$
C_{y_p}	$\frac{\partial (F_y/\bar{q}_b A)}{\partial (p\ell/2V)}$
C_{y_r}	$\frac{\partial (F_y/\bar{q}_b A)}{\partial (r\ell/2V)}$
C_{z_α}	$\frac{\partial (F_z/\bar{q}_b A)}{\partial \alpha}$
$C_{z_{\dot{\alpha}}}$	$\frac{\partial (F_z/\bar{q}_b A)}{\partial (\dot{\alpha} \ell/2V)}$
C_{z_u}	$\frac{\partial (F_z/\bar{q}_b A)}{\partial (u/V)}$
d	Body diameter
F_x	Body force along x-axis

F_y	Body force along y-axis
F_z	Body force along z-axis
f	Axis displacement from c.g. along x_1 -direction
G_{ij}	Elements of axis displacement matrix
g	Axis displacement from c.g. along y_1 -direction
H	Angular momentum
h	Axis displacement from c.g. along z_1 -direction
I_{ij}	Inertia components
ℓ	Reference length
L	Rolling Moment
M	Pitching Moment
$\mathcal{M}$	Mach number
$\mathfrak{M}$	Moment cross product term
N	Yawing moment
P	Angular velocity vector
p	Roll rate
q	Pitch rate
$\bar{q}$	Dynamic pressure, $\frac{1}{2}\rho V_\infty^2$
r	Yaw rate
t	Time
U	Translational velocity vector
u	Velocity in x-direction
V_∞	Total velocity, $\sqrt{u^2+v^2+w^2}$
$\mathcal{V}_i$	Velocity cross product term
v	Velocity in y-direction
w	Velocity in z-direction

X	Force vector
$\bar{X}$	Axis displacement vector
x	Coordinate
y	Coordinate
z	Coordinate
α	Angle of attack, $\sin^{-1}(w/V_\infty)$
$\bar{\alpha}$	Total angle of attack, $\sin^{-1}(\sqrt{v^2+w^2}/V_\infty)$
β	Angle of sideslip, $\sin^{-1}(v/V_\infty)$
δ	Asymmetry yaw angle
ϵ	Asymmetry pitch angle
ϵ_{ijk}	Permutation tensor
ρ	Mass density of atmosphere
σ	Asymmetry roll angle
Ψ	Projectile yaw angle
Θ	Projectile pitch angle
Φ	Projectile roll angle
ω	Angular velocity vector
λ	Angle of attack function

Subscripts

$(\ )_o$	Reference condition
$(\ )_1$	c.g. axes
$(\ )_2$	other axes

I. INTRODUCTION

The use of a free flight ballistic range to determine the aerodynamic characteristics of various types of munitions has long been standard practice. However, the experimental and analytical methods employed have steadily improved through the years. In the past decade advances have been made which significantly expand the field of application and accuracy of ballistic range testing. The Aeroballistic Research Facility at Eglin AFB represents the current state of the art in ballistic ranges. The analytical data extraction techniques in use at this facility also represent the most recent trends in the determination of aerodynamic data by free flight measurement.

In spite of the excellence of the Eglin procedures there is evidence to suggest that improvements can be made. Excellent results are consistently achieved for symmetrical, spin stabilized projectiles. Some tests of finned bodies have also been quite successful. It appears, however, that for airplane-like projectiles and for projectiles with mass and/or geometrical asymmetries the results achieved are not consistent. This has led to the present study of the aerodynamic models now used in the data extraction programs.

This report documents the studies performed during the summer program. It also includes an outline for continued development of the aerodynamic and dynamic models incorporated in the coefficient extraction programs.

II. OBJECTIVES

The primary objectives of the present work are:

1. To review current data extraction methods and expose any deficiencies in dynamic and aerodynamic models.

2. To investigate new methods for modeling aerodynamic forces and moments.

3. To develop a systematic procedure whereby aerodynamic models can be generated and verified with special emphasis being placed on the dependence of the models on geometric and mass asymmetries.

III. METHODOLOGY REVIEW

The remarks made in this section will deal with the aerodynamic data extraction procedures employed by the Ballistics Branch at Eglin Air Force Base. Figure 1 is a flow chart of the data analysis process. A test item is first inspected, weighed and measured. The inertias and center of gravity are determined (station 1). The item is then fired in the ballistic range (station 2) where shadowgraphs in orthogonal planes and time-of-passage measurements are recorded at each of 50 stations along the flight path. These constitute the test data.

The test data are processed in the film reading laboratory (station 3) where precise translational position and angular orientation are determined from the shadowgraph pictures and then tabulated and digitized in terms of tunnel coordinates (x, y, z) and test item orientation angles (α, β, ϕ). These digitized data are then utilized in the numerical coefficient extraction programs (station 5), which are run from a remote terminal at the Aeroballistic Research Facility. A detailed description of the test facility and instrumentation may be found in Reference 1.

The aerodynamic coefficient extraction procedures currently being used are described in References 2 and 3. Both programs use analytical models patterned after those of Reference 4. One program (Reference 2) is designed to handle configurations with at least three planes of symmetry while the other can supposedly deal with analysis of more complex shapes. Both programs make use of the Chapman-Kirk least-squares extraction technique described in Reference 5. This method, which minimizes the error between test data and analytically predicted results by a least-squares method, requires that the equations of motion be solved in an uncoupled form because the units and magnitudes of linear and angular displacements are different, and data quality may be inconsistent.

The present methods do a very good job of extracting aerodynamic coefficients for conventional symmetrical shapes such as artillery shells and bullets. Apparently the aerodynamic models for these shapes are adequate, and the dynamic coupling is very small. For symmetrical shapes with fins, however, and for unsymmetrical shapes the methods are less consistent. To date insufficient testing and extraction with the unsymmetrical program has been done to determine why the inconsistencies occur. In fact this is the motivation behind the involvement of the author in the present studies.

Careful review of the unsymmetrical program described in Reference 3 has revealed certain inconsistencies in the analytical models. Some sign errors which appear in the transformation equations in the report have, fortunately, been corrected in the computer program. However, the sign on the magnus force term, $Cy_{p\alpha}$ still appears to be different from normal ballistic convention. In addition the method of accounting for a

mass asymmetry appears to be very specialized in both the report and the program. In fact, this induced the author to derive the generalized equations described in part IV of this report. The aerodynamic models are in the conventional form based on Taylor series expansions of the forces and moments. The various forces and moments are represented as even (drag, damping) or odd (lift, pitching moment, etc) depending on the normal behavior of these terms. While there are terms to allow for non-linearities, the present models do not contain terms which might occur for non-symmetrical configurations.

Other possible extraction methods which are currently under study were reviewed. The maximum likelihood method described in References 6 and 7 is a logical extension of the method now being used. This method makes use of statistical data pertaining to experimental accuracy to alleviate the problems caused in the least-squares method by data quality and magnitude. This makes possible the simultaneous solution of the translational and rotational equations of motion.

Another approach to coefficient extraction is the extended Kalman filter procedure outlined in References 8 and 9. The extended filter can estimate the states of non-linear systems in the presence of noise, and it incorporates prior information on values for various coefficients without compromising the accuracy of the extraction process. This technique and the maximum likelihood methods have still not been thoroughly tested, but they are promising and do appear to be methods of dealing with highly coupled motions which the least-squares methods may not be able to handle.

During the reviewing process, the author gained an appreciation for the methods of experiment and analysis associated with ballistic testing. In addition, it became apparent that the utility of an outstanding test range (the Aeroballistic Research Facility), or any ballistic range, for that matter, could be enhanced by further refinement of aerodynamic models and extraction techniques. This knowledge served as a strong motivation for continued studies of aerodynamic models.

IV. DEVELOPMENT OF ASYMMETRY TRANSFORMATIONS*

A general trajectory analysis program should have the capability of dealing with a variety of shapes. This implies that provisions should be made for mass and geometric asymmetries and for differences in aerodynamic references axes. It is important in evaluating the aerodynamic forces and moments of a configuration that consistent data be used. This means that data from wind tunnel, free flight ballistic range and flight test must be provided in a common way in order that they can be meaningfully compared.

In ballistic analysis the equations of motion are conveniently written with respect to the center of gravity of the test item. If the cg is offset from the aerodynamic reference axes (e.g. a body centerline), the aerodynamic

* Detailed equations are shown in the Appendix.

coefficients extracted will be referenced to the c.g. and may not be comparable to aerodynamic data from other sources which are referred to the aerodynamic axes. The requirement for consistent coefficient values necessitates the development of a transformation procedure. Such a procedure is described in the following paragraphs. It should be pointed out that the transformation as shown is very general. For most cases only one or two offsets would be required thereby simplifying the transformation procedure. In addition for some coefficients the transformations may show negligible alterations in numerical values.

Figure 2 illustrates the axes transformation under consideration. Axes (x_1, y_1, z_1) represent a set positioned at the center of gravity. Axes(x_2, y_2, z_2) are another set positioned at another location, usually on the aerodynamic reference point or on a plane of symmetry. The linear displacements (f, g, h) and Euler angles (δ, ϵ, σ) define the orientation of (x_2, y_2, z_2) with respect to (x_1, y_1, z_1).

Coordinate Transformations - The axes coordinates can be transformed in the usual way by first translating the origin and then rotating first through δ (about z_1), then through ϵ (about y_1') and finally through σ (about x_2). The resulting transformations are:

$$\{X_2\} = [B]\{X_1 - \bar{X}\} \tag{1}$$

and

$$\{X_1\} = [B]^T\{X_2\} + \{\bar{X}\} \tag{1a}$$

Inertia Transformations - The inertias about axes (x_2, y_2, z_2) in terms of the inertias about axes (x_1, y_1, z_1) can be computed in the following way (using Ixx and Iyz as typical components):

$$I_{xx_2} = \int (y_2^2 + z_2^2)\,dm \quad \text{and} \quad I_{yz_2} = \int y_2 z_2\,dm$$

But, according to the coordinate transformations

$$y_2 = b_{21}x_1 + b_{22}y_1 + b_{23}z_1 \quad \text{and} \quad z_2 = b_{31}x_1 + b_{32}y_1 + b_{33}z_1$$

so that the integrals can be evaluated in terms of (x_1, y_1, z_1). The results of the transformations are:

$$\{I_2\} = [C]\{I_1\} + \{G\} \tag{2}$$

Translational Velocity Transformations - When the origins of two axes systems are not coincident a set of translational velocity components is induced on one set by virtue of the angular rates on the other. The components of these induced velocities are given by

$$V_i = \epsilon_{ikj}\,\bar{x}_k\,\omega_j$$

where $\bar{X} = f\bar{\imath} + g\bar{\jmath} + h\bar{k}$

$\omega = p_1\bar{\imath} + q_1\bar{\jmath} + r_1\bar{k}$

Hence, the velocity transformations become

$$\{U_2\} = [B]\{U_1 - \mathcal{V}\} \tag{3}$$

$$\{U_1\} = [B]^T\{U_2\} + \{\mathcal{V}'\} \tag{3a}$$

Rotational Rate Transformations - These quantities transform in the same way as the coordinates in rotation, and they are unaffected by translation. Hence, they can be written

$$\{P_2\} = [B]\{P_1\} \tag{4}$$

$$\{P_1\} = [B]^T\{P_2\} \tag{4a}$$

Force Transformations - The forces transform in the same manner as do the rotational rates. Therefore

$$\{F_2\} = [B]\{F_1\} \tag{5}$$

$$\{F_1\} = [B]^T\{F_2\} \tag{5a}$$

Moment Transformations - The moments transform in the same way as do the translational velocities. The moments induced by the translational displacement of a force system are given by

$$\mathcal{M}_i = \epsilon_{ikj}\,\bar{X}_k\,F_j$$

The moment transformations become

$$\{M_2\} = [B]\{M_1 - \mathcal{M}\} \tag{6}$$

$$\{M_1\} = [B]^T\{M_2\} + \{\mathcal{M}'\} \tag{6a}$$

Aerodynamic Coefficient Transformations - When reference axes are displaced, the aerodynamic coefficients defined in terms of one axes set must be transformed to those of another axes set. It is a very tedious process to derive the most general forms of these coefficients and, performing this task for all coefficients is beyond the scope of this work. However, the results for four typical coefficients are given in the Appendix. A typical derivation follows.

Consider the case where characteristics known with respect to axes (x_2, y_2, z_2); an aerodynamic reference axes set, are to be transformed to the center of gravity axes set (x_1, y_1, z_1). Typically,

$$F_{x_{\alpha_1}} = \frac{\partial F_{x_1}}{\partial \alpha_1} = \frac{\partial F_{x_1}}{\partial w_1} \frac{\partial w_1}{\partial \alpha_1}$$

$$= \frac{\partial}{\partial w_1} \left(b_{11} F_{x_2} + b_{21} F_{y_2} + b_{31} F_{z_2} \right) \frac{\partial w_1}{\partial \alpha_1} \qquad (7)$$

But, typically $$\frac{\partial F_{x_2}}{\partial w_1} = \frac{\partial F_{x_2}}{\partial w_2} \frac{\partial w_2}{\partial w_1} + \frac{\partial F_{x_2}}{\partial v_2} \frac{\partial v_2}{\partial w_1} + \cdots + \frac{\partial F_{x_2}}{\partial q_2} \frac{\partial q_2}{\partial w_1} + \cdots \qquad (7a)$$

From (3) and (4) these derivatives can be evaluated to be

$$\frac{\partial w_2}{\partial w_1} = b_{33} \; ; \; \frac{\partial v_2}{\partial w_1} = b_{23} \; ; \; \frac{\partial P_i}{\partial w_1} = 0 \qquad (7b)$$

Also $$\sin \alpha_1 = w_1 / V_\infty \; ; \quad \sin \alpha_2 = w_2 / V_\infty$$

Thus $$\frac{\partial w_1}{\partial \alpha_1} = V_\infty \cos \alpha_1 \; ; \quad \frac{\partial \alpha_2}{\partial w_2} = \frac{1}{V_\infty \cos \alpha_2} \qquad (7c)$$

Note also that $$\frac{\partial F_{x_2}}{\partial w_2} = \frac{\partial F_{x_2}}{\partial \alpha_2} \frac{\partial \alpha_2}{\partial w_2} = \frac{1}{V_\infty \cos \alpha_2} \frac{\partial F_{x_2}}{\partial \alpha_2} \qquad (7d)$$

Now, let $$C_{x_\alpha} \equiv \frac{\partial}{\partial \alpha} \left[\frac{F_x}{(\frac{1}{2} \rho V_\infty^2) A} \right] \qquad (7e)$$

Substitution of equations (7a) through (7e) back into equation (7) yields, for a case where certain derivatives are zero (e.g. $\partial F_{y_2} / \partial w_1 = 0$),

$$C_{x_{\alpha_1}} = \lambda_1 \left[\cos\delta \cos^2\epsilon \cos\sigma \, C_{x_{\alpha_2}} + \tfrac{1}{2} \left(\cos\delta \sin 2\epsilon \cos^2\sigma + \sin\delta \cos\epsilon \sin 2\sigma \right) C_{z_{\alpha_2}} \right]$$

$$- \lambda_2 \left[\tfrac{1}{2} \cos\delta \sin 2\epsilon \, C_{x_{u_2}} + \left(\cos\delta \sin^2\epsilon \cos\sigma + \sin\delta \sin\epsilon \sin\sigma \right) C_{z_{u_2}} \right] \qquad (7f)$$

where $$\lambda_1 = \frac{\cos \alpha_1}{\cos \alpha_2} \; ; \quad \lambda_2 = \cos \alpha_1$$

It should be understood that (7f) is not the most general form for C_{x_α}, nor is it necessarily the most practical. Further work will be done in this area to compile a set of practical working equations for use by the ballistician.

It is of interest to note that if only a single-plane rotational displacement is considered (i.e., $\bar{X}=0$ and $\delta=\sigma=0$), equation (7f) reduces to

$$C_{x_{d_1}} = \cos^2\epsilon\, C_{x_{\alpha_2}} + \sin\epsilon\cos\epsilon\,(C_{z_{\alpha_2}} - C_{x_{u_2}}) - \sin^2\epsilon\, C_{z_{u_2}}$$

which is identical to the equation transformed in Appendix 3 of Reference 10.

V. AERODYNAMIC MODEL ANALYSIS

The aerodynamic model is the framework within which the entire extraction process must take place. In light of the comments in Section IV of this report, the equations of motion for a projectile in free flight may be written

$$\dot{\overline{mV}} = \sum \bar{F} \tag{8}$$

$$\dot{\bar{H}} = \sum \bar{M} \tag{9}$$

Thus, there are three equations of translation [(8)] and three equations of rotation [(9)]. The forces and moments in these equations include aerodynamic terms. It is the specific forms of these aerodynamic terms that are of interest here (i.e., aerodynamic force and moment models).

The conventional method of dealing with forces and moments in trajectory and stability analysis is to expand them in a multi-variable Taylor series about some reference values of the independent variables. As examples, the axial force, Fx, and pitching moment, M, would be expanded as follows:

$$F_x = F_{x_o} + \frac{\partial F_x}{\partial \alpha}(\alpha-\alpha_o) + \frac{\partial F_x}{\partial u}(u-u_o) + \frac{\partial F_x}{\partial q}(q-q_o) + \cdots$$
$$+ \frac{1}{2!}\left[\frac{\partial^2 F_x}{\partial \alpha^2}(\alpha-\alpha_o)^2 + \frac{\partial^2 F_x}{\partial u^2}(u-u_o)^2 + \frac{\partial^2 F_x}{\partial q^2}(q-q_o)^2 + \cdots\right] + \cdots \tag{10}$$

$$M = M_o + \frac{\partial M}{\partial \alpha}(\alpha-\alpha_o) + \frac{\partial M}{\partial u}(u-u_o) + \frac{\partial M}{\partial q}(q-q_o) + \cdots$$
$$+ \frac{1}{2!}\left[\frac{\partial^2 M}{\partial \alpha^2}(\alpha-\alpha_o)^2 + \frac{\partial^2 M}{\partial u^2}(u-u_o)^2 + \frac{\partial^2 M}{\partial q^2}(q-q_o)^2 + \cdots\right] + \cdots \tag{11}$$

where the reference conditions are $(\quad)_o$, and the partial derivatives are evaluated at these reference conditions. Of course, the expansion can include terms of order higher than 2. In fact the axial force, Fx, and pitching moment, M, were selected as examples to point out typical characteristics in aerodynamic models presently being used.

The so-called "unsymmetrical program" described in Reference 3 uses the following models for axial force and pitching moment:

$$C_X = C_{X_o} + C_{X_{\alpha 2}}\left(\frac{w}{V_\infty}\right)^2 + C_{X_{\beta 2}}\left(\frac{v}{V_\infty}\right)^2 + C_{X_M}(M - M_{REF})$$

$$C_m = C_{m_o} + C_{m_\alpha}\left(\frac{w}{V_\infty}\right) + C_{m_{\alpha 3}}\left(\frac{w}{V_\infty}\right)^3 + \left(\frac{qd}{2V_\infty}\right)\left[C_{m_q} + C_{m_{q2}}\left(\frac{w}{V_\infty}\right)^2\right]$$

$$+ C_{m_{p\beta}}\left(\frac{v}{V_\infty}\right)\left(\frac{pd}{2V_\infty}\right) + \sin^2\bar{\alpha}\sin(n\phi')\left[C_{m_{\phi\alpha}}\left(\frac{w}{V_\infty}\right) + C_{n_{\phi\alpha}}\left(\frac{v}{V_\infty}\right)\right]$$

where $C_X = \dfrac{F_X}{(\frac{1}{2}\rho V_\infty^2 A)}$ $\qquad$ $C_m = \dfrac{M}{(\frac{1}{2}\rho V_\infty^2) Ad}$

and $\phi' = \tan^{-1}\left(\dfrac{v}{w}\right) - \zeta$ $\qquad$ ζ = phase angle of fins

Several interesting things are apparent from these expansions. These are:

1. The axial force and a damping term, being primarily dissipative processes, are represented as even functions of the flow angle because they should be of the same sign regardless of the flow direction.

2. The pitching moment is represented as an odd function of flow angle since, in normal cases, the slope of the moment-inclination angle curve does not change sign.

3. The roll orientation term, which would be particularly important for certain winged or finned configurations, is actually a cubic (odd function) in terms of flow inclination angle, though it is represented as the product of $\sin^2\bar{\alpha}$ and w/V_∞ (or v/V_∞) because the net flow actually approaches the body along $\bar{\alpha}$, but the contribution from each component of this flow must be given in terms of w/V_∞ (or v/V_∞). Of course, the actual body orientation ($n\phi'$) is represented only linearly in $\sin(n\phi')$.

Figures 3 and 4 show the axial force and pitching moment for a typical ballistic configuration. The models fit the data quite well, particularly at low and moderate angles of attack. Incidentally, the extraction process,

being a least-squares curve fitting operation, actually determines the coefficients by minimizing the errors between experimental data and the analytical model all along the range of angle of attack. These coefficients are not true Taylor series coefficients, then, but it is really immaterial as far as the results.

The good model fit shown in Figures 3 and 4 is not necessarily representative of all coefficients for that configuration nor of any coefficients for some other configuration. In fact, apparently the more "unconventional" is the configuration, the more unreliable is the model likely to be. Notice that for the results shown in Figures 5 and 6, the model fit is poor at the higher angles of attack even though the configuration is a conventional ogive-cylinder with tail fins. There have been some cases for which inclusion of non-linear terms in one force or moment have affected the coefficients of some other force or moment. This, then, is the essence of the problem being considered - to develop aerodynamic models for inclusion in the extraction process which will be reliable for all configurations tested in a ballistic range.

The first alternative is to use series models with more terms. For instance, for a configuration which is truly unsymmetrical, it is possible that an odd or even function formulation is not appropriate. Thus, for certain forces or moments both odd and even powers of the independent variables may be required in the model. A typical variation is shown in Figure 7.

Other analytical models may also be feasible. In dynamic simulation table look up routines are often used. Though such a scheme would not be compatible with the present extraction algorithm, perhaps a related method can be found. In fact one of the major long term studies which is suggested is a detailed look at the classes of aerodynamic models which are consistent with both the extraction process and the actual aerodynamic behavior of ballistic munitions.

In order to gain an appreciation for the aerodynamic behavior of some ballistic shapes a computer search at the Eglin AFB Technical Library was performed. Key words ASYMMETRIC BODIES - WIND TUNNEL TESTS yielded six references. In addition several more references were found from other sources. Because of the time limitations imposed by the summer schedule, a thorough search was not made. Nevertheless, a few typical plots were made and a power series of the type now being used was matched to the data. It is seen that for the coefficient of Figure 8, the present model does very well, in spite of the fact that the configuration has an elliptical body with fins. Unfortunately, for the droop-nosed configuration of Figure 9, there is little chance of fitting the data with the present model.

To date specific models which will account for non-symmetrical effects have not been developed. However, a general research plan has been formulated and certain key elements in the modeling process have been identified. These are discussed in the next section.

VI. RESEARCH PLAN

The development of a comprehensive long range research plan for any problem is a formidable task. The present problem is no exception. However, the summer study effort has pointed out some paths to follow, and all that remains is to formulate the details associated with these paths. Some details can be provided now; others will have to evolve as the overall plan proceeds. All are dependent on the final objective to be achieved. As the author sees it, there are both short range and long range objectives. The immediate objective is to develop aerodynamic models with which confident predictions from aeroballistic data can be made for munitions in the present inventory and on the near horizon. The long range objective is to determine the limits, from an aerodynamic modeling viewpoint, of applicability of aeroballistic testing to determination of aerodynamic coefficients.

The proposed research plan involves four main parts. These are highlighted as follows:

1. Modification of existing programs - The existing dynamics programs are not restricted to munitions with several planes of symmetry. However, more generality should be added to better account for mass, geometry and aerodynamic asymmetries. In this phase of the research the aerodynamic models would be left intact, but the dynamics would be modified to permit the use of any axis system in the analysis. Provisions would be made, as described in Section IV for transforming physical and flow properties between axis systems. With the inclusion of such modifications the existing programs could then be run to determine the effects of these changes on accuracy. Standard 6-DOF simulations could be used to perform some of these tests.

2. Investigate existing extraction programs - There is no doubt that the aerodynamic model used in the extraction programs produces a bias on the results. There is also probably a configuration bias also. The quality of results depends on both the extraction program and the aerodynamic models. Therefore, before new models are sought some effort should be made to determine the compatibility of model types with the extraction process. Involved in this would be the determination of acceptable functional (or discrete) forms. Also, an effort should be made to find out the effects of uncoupling translational and rotational dynamics and possible ways to improve the procedure if these effects are severe for unsymmetrical shapes. The existing extraction process is sensitive to the order in which coefficients are sought. While this may be entirely due to the fact that local rather than global minima are selected, some studies should be made of these effects.

3. The literature should be searched for appropriate wind tunnel data from which characteristic aerodynamic models can be constructed. It is possible that specific "class" characteristics could be detected which would permit the generation of selected configuration-oriented models which could be used in modular form as various configurations are tested. While these models would not be general, they would hold the promise of providing good results for specific classes of munitions. Once models are determined, they should be tested with a synthetic six-DOF simulation program and by comparison with existing test data.

4. It will be absolutely necessary to plan a test program which will highlight various phases of the modeling problem. Special models should be designed which emphasize various aspects of the asymmetrical munitions problem, whether they be related to mass or geometry. Such tests will be used to verify new model forms, detect inconsistencies in the extraction process.

It is obvious from the preceding paragraphs that an extensive research program is proposed. However, it is imperative that such an outstanding facility as the Aeroballistic Research Facility be used to the fullest extent and that the results forthcoming from it be used with confidence. It is believed that this program will aid in achieving those goals.

VII. SPECIAL SEMINARS AND BRIEFINGS

During the course of the summer program the author has recieved a large amount of information via personal and group briefings and seminars. Initially, a very excellent tour and summary of operations of the Aeroballistic Research Facility (ARF) was provided by Lt Demo Galanos. A description of the data processing procedures, computer terminal operations and preliminary ballistic testing were provided by Mr. Kenneth West. Messrs. Kenneth Cobb and Gerald Winchenbach each provided thorough briefings of the theoretical background and computational details of the extraction programs. Mr. Cobb also presented a formal seminar on the dynamics of an impulse corrected, laser guided rocket (Reference 13) which provided considerable insight into the theoretical methods applicable in rocket trajectory analysis.

Two formal seminars were given by staff members of the Analytical Sciences Corporation (TASC). Dr. Charles Brown talked on applied optimal estimation methods. His talk and subsequent discussions with him were very helpful in developing an appreciation for stochastic methods. In particular, his discussion of Kalman filters was useful in pointing out the application of these methods in aerodynamic coefficient estimation.

Mr. Jim Kain's topic of discussion was aerodynamic coefficient estimation from flight test data. Much of what he said has direct application to the author's modeling study, though the flight test data extraction is more complex than the aeroballistic case becasue of instrumentation differences and because of the uncontrolled environment. One interesting facet of this seminar was the introduction of the concept of structure identification. This implies the formulation of an extraction algorithm which would select its on structure (aerodynamic model) to fit any particular case. The so-called "structure" would be in the general series format, but terms of the series would be selected by the extraction program. This is being pursued by Mr. Phil Fiske of TASC. All of the work being done in this area by this contractor is currently using the extended Kalman filter technology. The results have been encouraging.

VIII. CONCLUSIONS AND RECOMMENDATIONS

During the ten-weeks' duration of the Summer Research Program the author has become familiar with the experimental methods and analytical tools currently employed in the aeroballistic coefficient extraction procedure. Some of the problems associated with this procedure have also come to light. As a result of personal studies and numerous discussions with branch and division personnel it is concluded that better aerodynamic models are required not only to better determine the aerodynamic characteristics of existing munitions, but to extend the analysis capabilities of the ballistic range to more unconventional shapes.

In order to begin the evolution of more appropriate aerodynamic models it is recommended that the following steps be taken:

1. Modify existing dynamics programs to better account for mass asymmetries and to permit transformation and computation of aerodynamic coefficients for any reference.

2. Investigate existing extraction programs to determine model and configuration bias effects and to determine compatibility of various model forms.

3. Using experimental data for a variety of ballistic configurations generate possible aerodynamic models.

4. Plan and carry out a testing program devoted strictly to assessing extraction program performance and aerodynamic model structure.

REFERENCES

1. Winchenbach, G. L., Judith S. Kleist, D. G. Galanos, and Goldman E. Parrish, "Description and Capabilities of the Aeroballistic Research Facility", presented at the Third Annual Automatic Cannon Caliber Symposium, Frankford Arsenal, Philadelphia, Pennsylvania, September 28, 1976.

2. Whyte, Robert H. and Wayne H. Hathaway, "Aeroballistic Range Data Reduction Technique Utilizing Numerical Integration", AFATL-TR-74-41, February 1974.

3. Hathaway, Wayne H. and Robert H. Whyte, "Aeroballistic Range Data Analysis for Unsymmetrical Configurations", AFATL-TR-76-109, September, 1976.

4. Brunk, James E., "Users' Manual: Extended Capability Magnus Rotor and Ballistic Body 6-DOF Trajectory Program", AFATL-TR-70-40, May, 1970.

5. Chapman, Gary T. and Donn B. Kirk, "A Method for Extracting Aerodynamic Coefficients from Test Data", AIAA Journal, Vol 8, No. 4, April, 1970

6. Grove, Randall D., Roland L. Bowles and Stanley C. Mayhew, "A Procedure for Estimating Stability and Control Parameters from Flight Test Data by Using Maximum Likelihood Methods Employing a Real-Time Digital System", NASA TN D-6735, May, 1972

7. Hathaway, Wayne H., "Analysis of the Free Flight Aerodynamics of Nonsymmetric Bodies from Ballistic Spark Range Data", MS Thesis, University of Vermont, February, 1976.

8. Daniel, Donald C., "An Analysis of Methods for Extracting Aerodynamic Coefficients from Test Data", AFATL-TR-73-32, February, 1973.

9. Brown, Charles M., Jr., "An Extended Kalman Filter for Estimating Aerodynamic Coefficients", AFATL-TR-76-158, December, 1976.

10. Babister, A. W., Aircraft Stability and Control, Pergamon Press, New York, 1961.

11. Smith, Dale K., "Aerodynamic Characteristics of Three Maneuvering Air-to-Air Missile Models at Mach Numbers from 0.5 to 1.6", AFATL TR 77-3, June, 1977

12. Scallion, William I. and Michael D. Cannon, "The Low Speed Static Longitudinal and Lateral Characteristics of a Delta Wing Model with Fixed and Free-Floating Canard Surfaces", NACA TM X-120, October, 1959.

13. Cobb, Kenneth K., "Dynamic Design Considerations for an Impulse Corrected, Laser Guided Rocket", presented at the AIAA Flight Mechanics Conference, Hollywood, Florida, August, 1977.

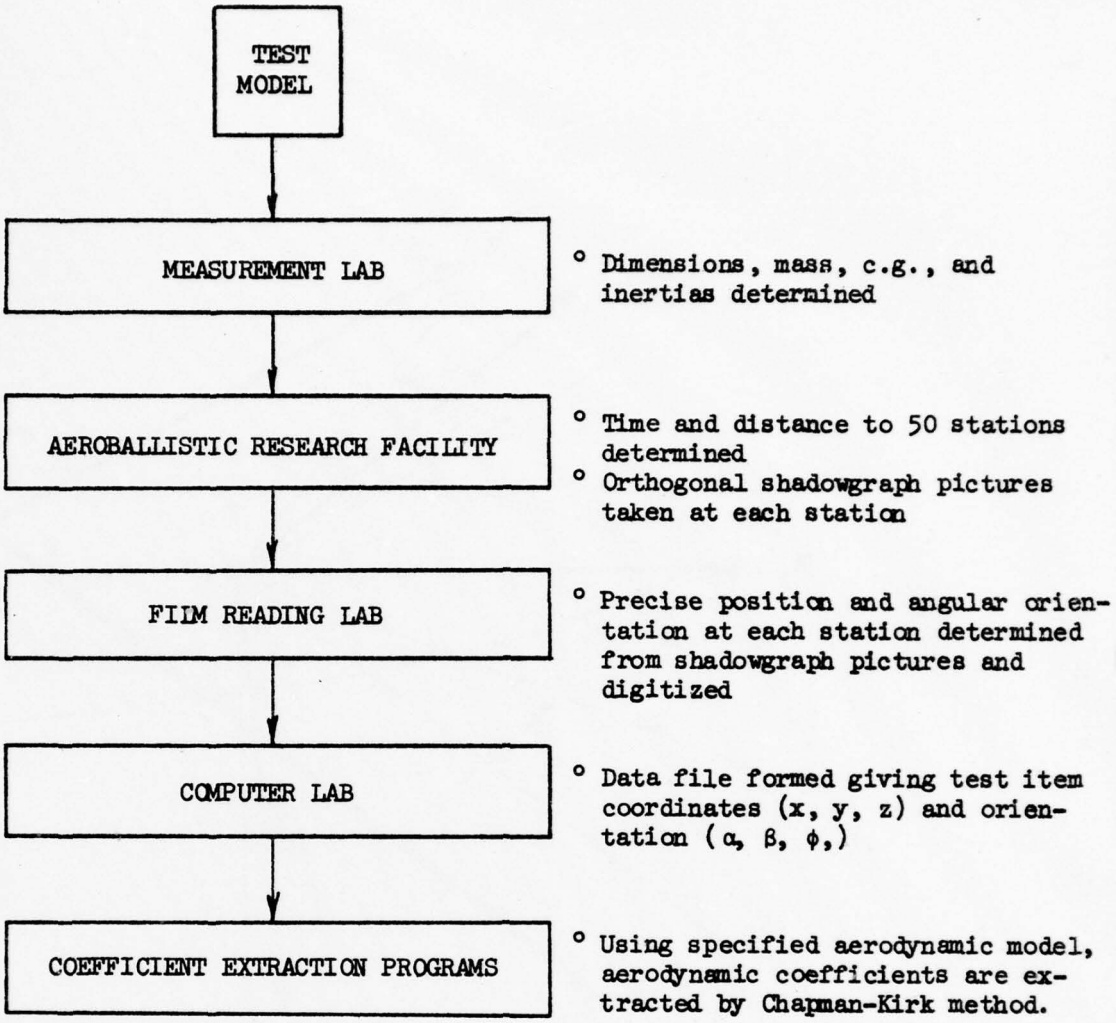


FIGURE 1. BALLISTIC RANGE COEFFICIENT EXTRACTION PROCESS

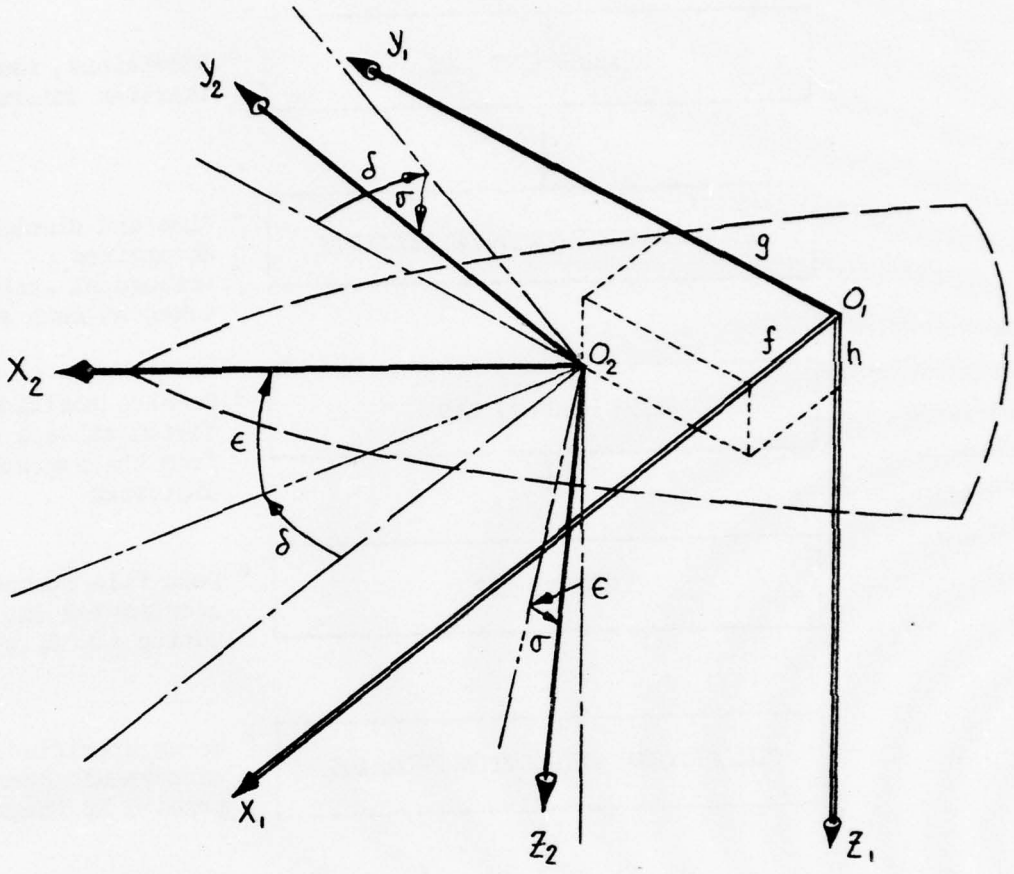


FIGURE 2. TRANSFORMATION NOMENCLATURE

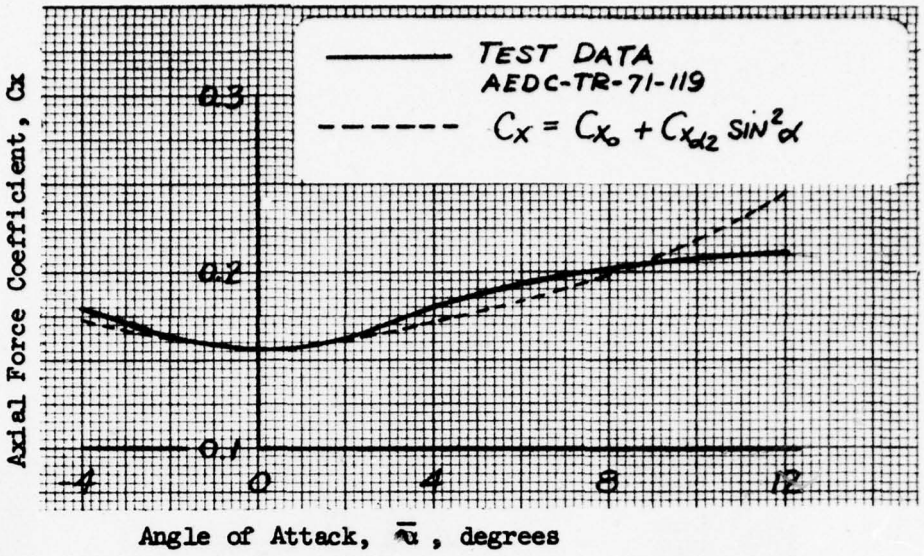


FIGURE 3. AXIAL FORCE COEFFICIENT FOR 5 CALIBER SPINNER PROJECTILE

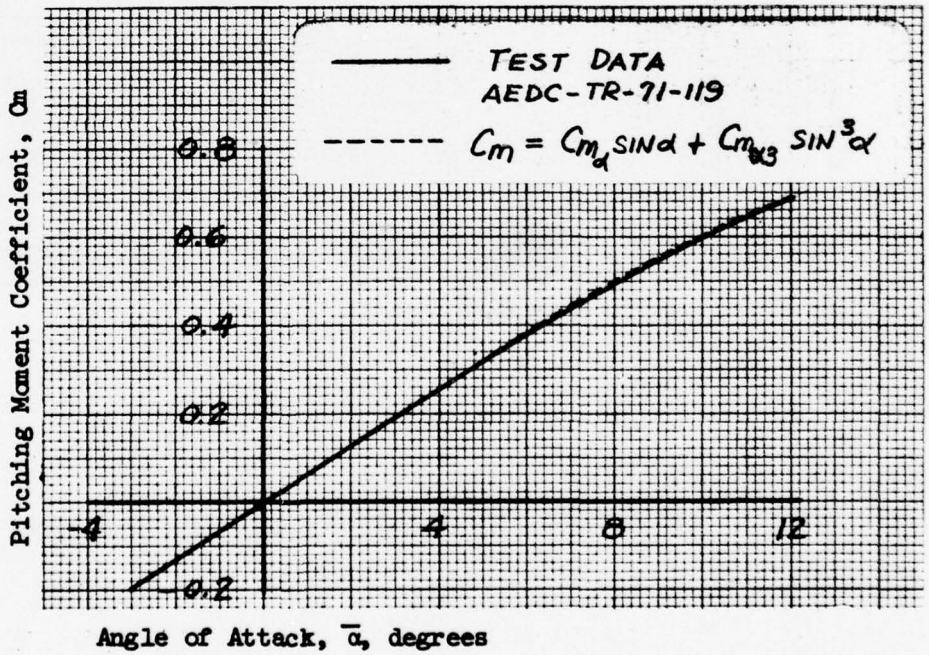


FIGURE 4. PITCHING MOMENT COEFFICIENT FOR 5 CALIBER SPINNER PROJECTILE

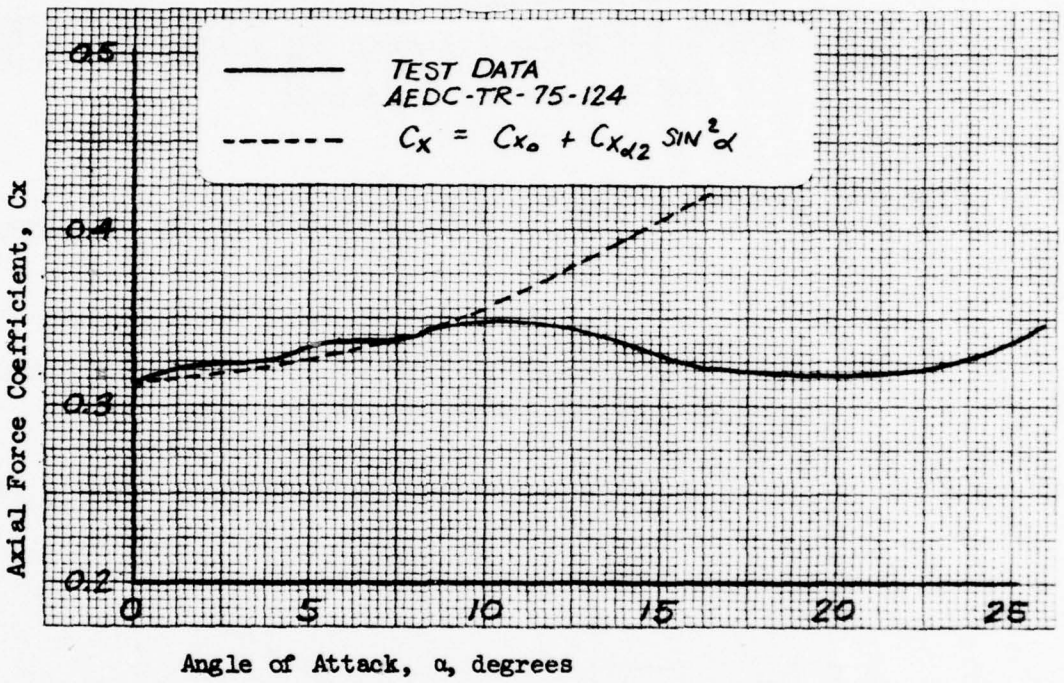


FIGURE 5. AXIAL FORCE COEFFICIENT FOR OGIVE-CYLINDER WITH TAIL FINS

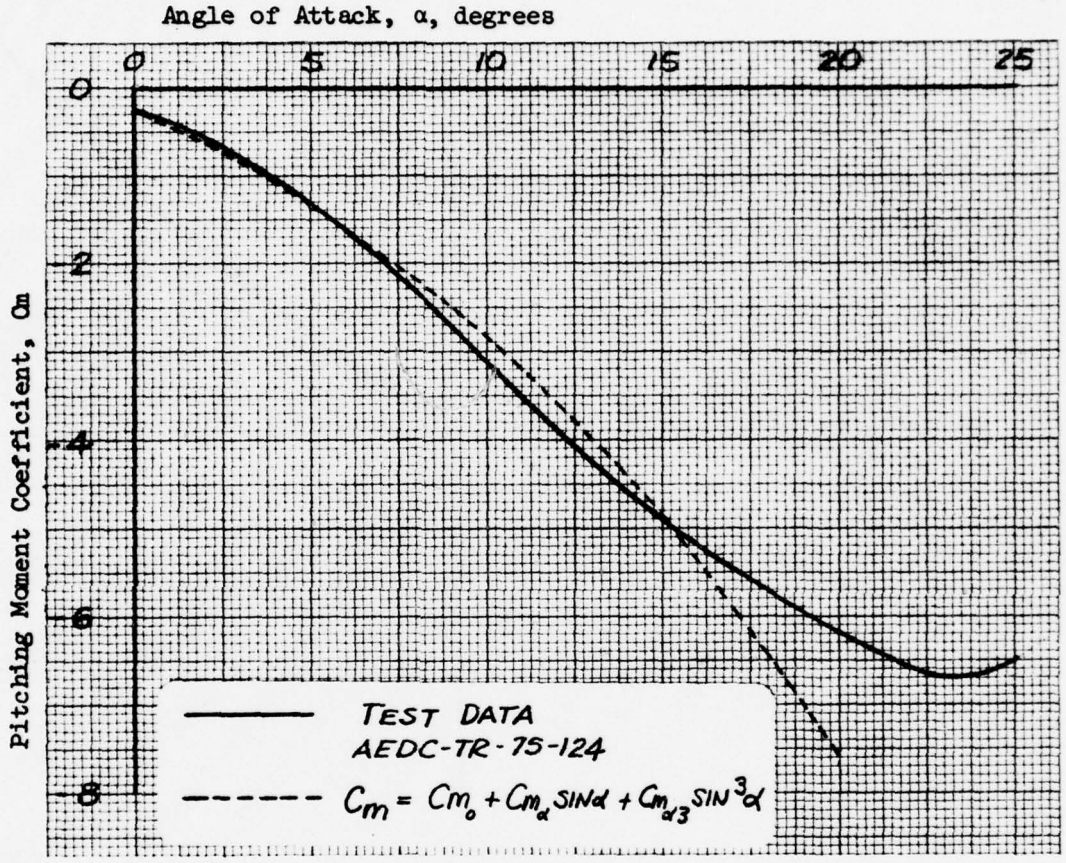


FIGURE 6. PITCHING MOMENT COEFFICIENT FOR OGIVE-CYLINDER WITH TAIL FINS

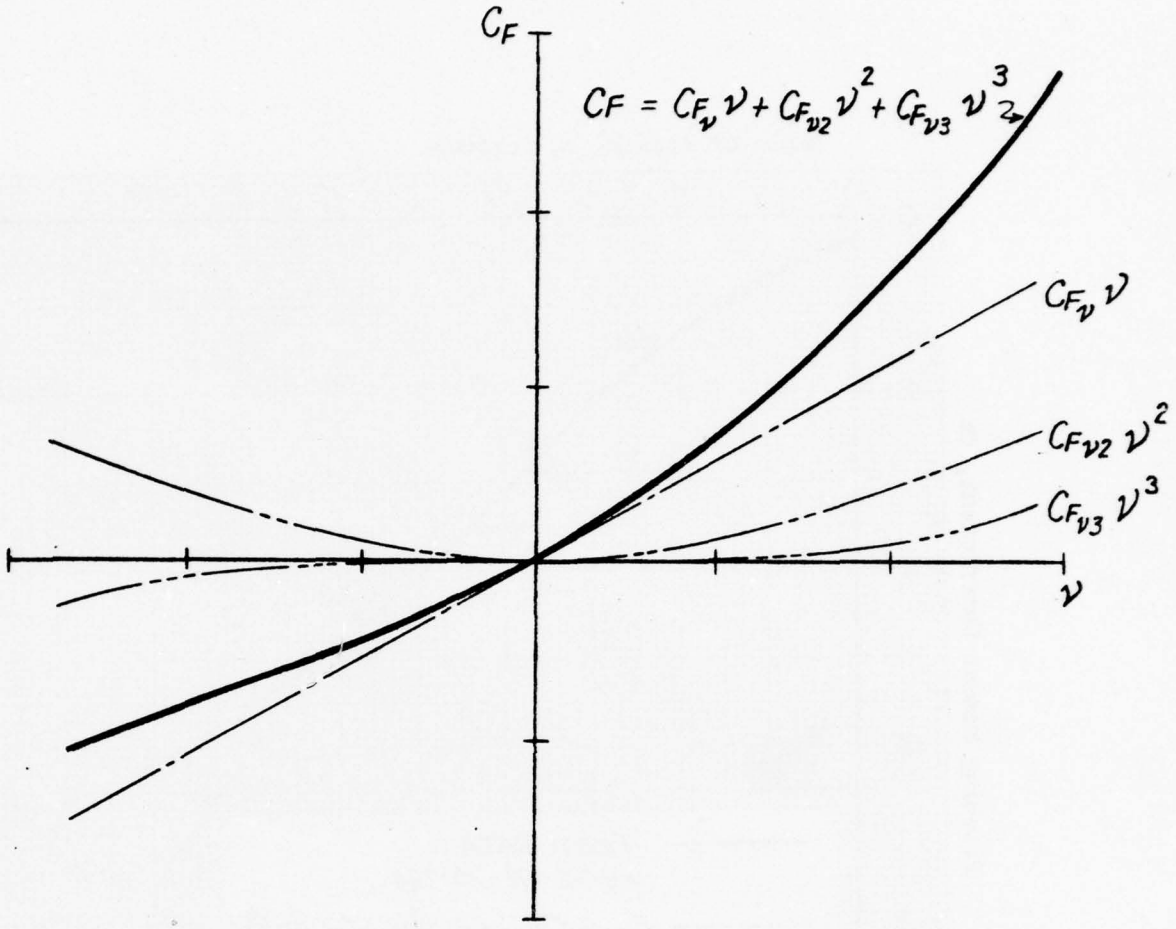


FIGURE 7. POSSIBLE COEFFICIENT VARIATION

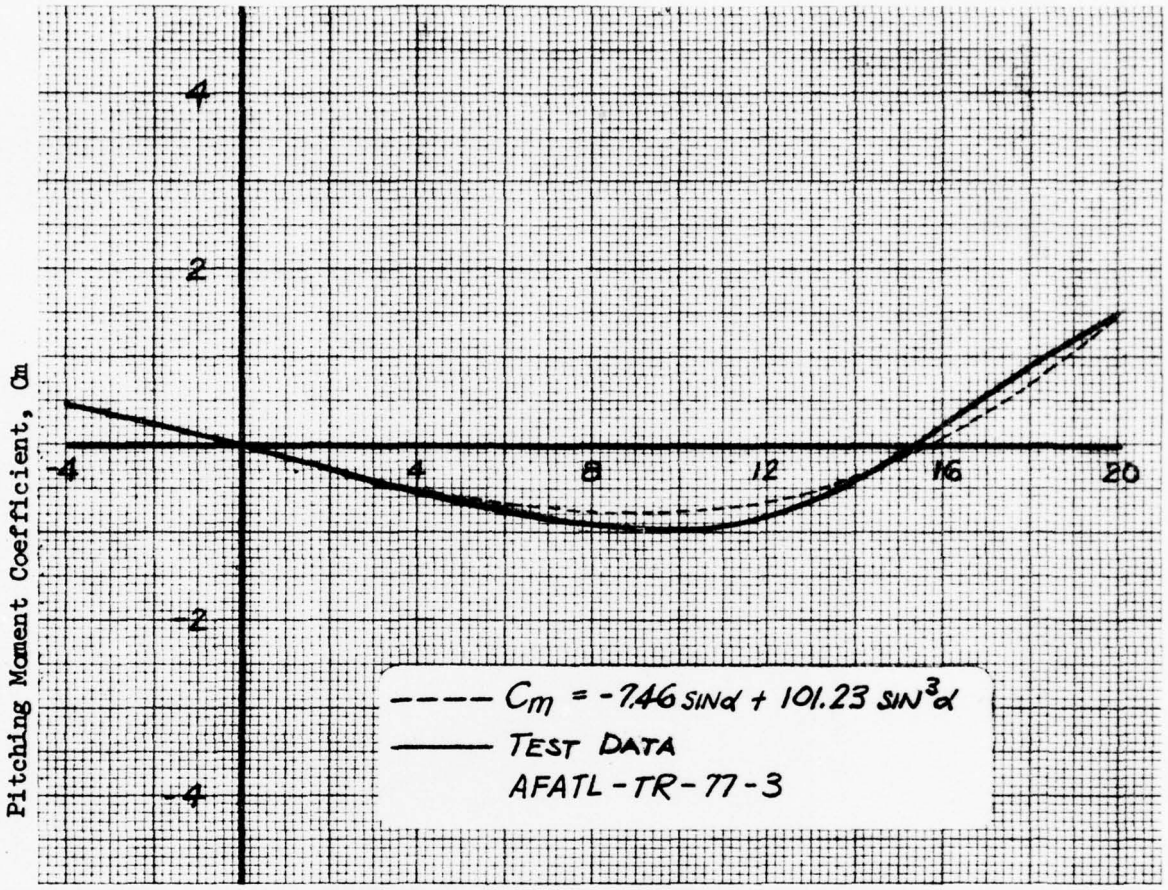


FIGURE 8. PITCHING MOMENT FOR FINNED, ELLIPTICAL BODY

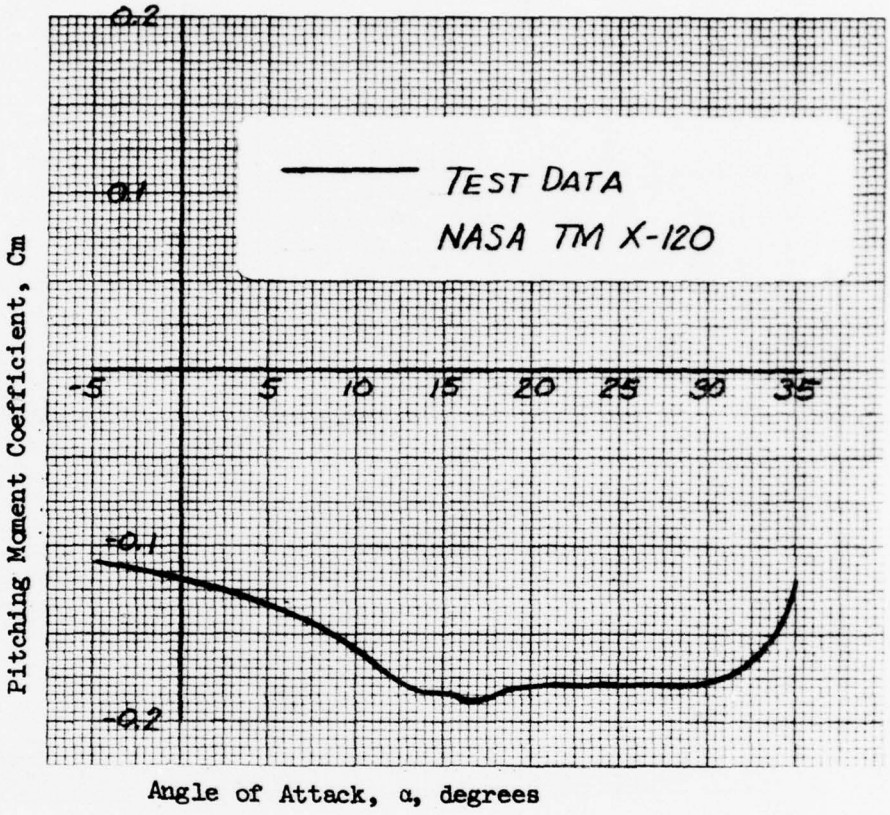


FIGURE 9. PITCHING MOMENT COEFFICIENT FOR DROOP-NOSED AIRPLANE

APPENDIX

List of Equations

1. Matrix elements, b_{ij}

$$[B] = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \qquad (A\text{-}1)$$

where

$$\begin{aligned} b_{11} &= \cos\delta\cos\epsilon \\ b_{12} &= \sin\delta\cos\epsilon \\ b_{13} &= -\sin\epsilon \\ b_{21} &= \cos\delta\sin\epsilon\sin\sigma - \sin\delta\cos\sigma \\ b_{22} &= \sin\delta\sin\epsilon\sin\sigma + \cos\delta\cos\sigma \\ b_{23} &= \cos\epsilon\sin\sigma \\ b_{31} &= \cos\delta\sin\epsilon\cos\sigma + \sin\delta\sin\sigma \\ b_{32} &= \sin\delta\sin\epsilon\cos\sigma - \cos\delta\sin\sigma \\ b_{33} &= \cos\epsilon\cos\sigma \end{aligned}$$

Equation (A-1) is the matrix (or its transpose) which appears in equations (1) and (3) through (7) of Section IV.

2. Matrix elements, C_{ij}

The C matrix appears in the inertia transformation which is given by equation (2) as

$$\begin{Bmatrix} I_{xx_2} \\ I_{xy_2} \\ I_{xz_2} \\ I_{yy_2} \\ I_{yz_2} \\ I_{zz_2} \end{Bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix} \begin{Bmatrix} I_{xx_1} \\ I_{xy_1} \\ I_{xz_1} \\ I_{yy_1} \\ I_{yz_1} \\ I_{zz_1} \end{Bmatrix} + \begin{Bmatrix} G_{xx} \\ G_{xy} \\ G_{xz} \\ G_{yy} \\ G_{yz} \\ G_{zz} \end{Bmatrix} \qquad (A\text{-}2)$$

where

$$C_{11} = \cos^2\delta \cos^2\epsilon \qquad C_{21} = \cos\delta\cos\epsilon\,(\cos\delta\sin\epsilon\sin\sigma - \sin\delta\cos\sigma)$$

$$C_{12} = -\sin 2\delta \cos^2\epsilon \qquad C_{22} = \cos 2\delta \cos\epsilon \cos\sigma + \tfrac{1}{2}\sin 2\delta \sin 2\epsilon \sin\sigma$$

$$C_{13} = \cos\delta \sin 2\epsilon \qquad C_{23} = \cos\delta\cos 2\epsilon \sin\sigma + \sin\delta\sin\epsilon\cos\sigma$$

$$C_{14} = \sin^2\delta\cos^2\epsilon \qquad C_{24} = -\tfrac{1}{2}\left[(1+\cos^2\delta)\sin 2\epsilon\sin\sigma - \sin 2\delta\cos\epsilon\cos\sigma\right]$$

$$C_{15} = \sin\delta\sin 2\epsilon \qquad C_{25} = \sin\delta\cos 2\epsilon\sin\sigma - \cos\delta\sin\epsilon\cos\sigma$$

$$C_{16} = \sin^2\epsilon \qquad C_{26} = \sin\epsilon\cos\epsilon\sin\sigma$$

$$C_{41} = \sin^2\delta\cos^2\sigma + \cos^2\delta\sin^2\epsilon\sin^2\sigma - \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{42} = \sin 2\delta\,(\cos^2\sigma - \sin^2\epsilon\sin^2\sigma) - \cos 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{43} = \sin\delta\cos\epsilon\sin 2\sigma - \cos\delta\sin 2\epsilon\sin^2\sigma$$

$$C_{44} = \cos^2\delta\cos^2\sigma + \sin^2\delta\sin^2\epsilon\sin^2\sigma + \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{45} = \cos\delta\cos\epsilon\sin 2\sigma + \sin\delta\sin 2\epsilon\sin^2\sigma$$

$$C_{46} = \cos^2\epsilon\sin^2\sigma$$

$$C_{31} = -\tfrac{1}{2}(\cos^2\delta\sin 2\epsilon\cos\sigma + \sin 2\delta\cos\epsilon\sin\sigma)$$

$$C_{32} = \tfrac{1}{2}(\sin 2\delta\sin 2\epsilon\cos\sigma) - \cos 2\delta\cos\epsilon\sin\sigma$$

$$C_{33} = \cos\delta\cos 2\epsilon\cos\sigma - \sin\delta\sin\epsilon\sin\sigma$$

$$C_{34} = \sin\delta\cos\epsilon\,(\cos\delta\sin\sigma - \sin\delta\sin\epsilon\cos\sigma)$$

$$C_{35} = \sin\delta\cos 2\epsilon\cos\sigma + \cos\delta\sin\epsilon\sin\sigma$$

$$C_{36} = \sin\epsilon\cos\epsilon\cos\sigma$$

$$C_{51} = \tfrac{1}{2}\,(\sin 2\delta\sin\epsilon\cos 2\sigma - \cos^2\delta\sin^2\epsilon\sin 2\sigma)$$

$$C_{52} = \tfrac{1}{2}\sin 2\delta\,(\sin^2\epsilon+1)\sin 2\sigma + \sin\epsilon\,(\cos 2\delta\cos^2\sigma + \cos^2\delta\cos 2\sigma)$$

$$C_{53} = \tfrac{1}{2}\cos\delta\sin 2\epsilon\sin 2\sigma - \sin\delta\cos\epsilon\cos 2\sigma$$

$$C_{54} = \tfrac{1}{2}\left[\sin 2\sigma\,(1-\sin^2\delta\sin^2\epsilon) - \sin 2\delta\sin\epsilon\cos 2\sigma\right]$$

$$C_{55} = \tfrac{1}{2}\sin\delta\sin 2\epsilon\sin 2\sigma + \cos\delta\cos\epsilon\cos 2\sigma$$

$$C_{56} = -\tfrac{1}{2}\cos^2\epsilon\sin 2\sigma$$

$$C_{61} = \cos^2\delta\sin^2\epsilon\cos^2\sigma + \sin^2\delta\sin^2\sigma + \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{62} = \sin 2\delta\left[1 - \cos^2\sigma\,(\sin^2\epsilon+1)\right] + \cos 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{63} = -\sin\delta\cos\epsilon\sin 2\sigma - \cos\delta\sin 2\epsilon\cos^2\sigma$$

$$C_{64} = \sin^2\delta\sin^2\epsilon\cos^2\sigma + \cos^2\delta\sin^2\sigma - \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma$$

$$C_{65} = \cos\delta\cos\epsilon\sin 2\sigma - \sin\delta\sin 2\epsilon\cos^2\sigma$$

$$C_{66} = \cos^2\epsilon\cos^2\sigma$$

and also,

$$G_{xx} = m\left[(1-\cos^2\delta\cos^2\epsilon)f^2 + (1-\sin^2\delta\cos^2\epsilon)g^2 + \cos^2\epsilon\, h^2 - fg(\sin 2\delta\cos^2\epsilon) + gh(\sin\delta\sin 2\epsilon) + fh(\cos\delta\sin 2\epsilon)\right]$$

$$G_{xy} = m\left[(\tfrac{1}{2}\cos^2\delta\sin 2\epsilon\sin\sigma - \tfrac{1}{2}\sin 2\delta\cos\epsilon\cos\sigma)f^2 + (\tfrac{1}{2}\sin^2\delta\sin 2\epsilon\sin\sigma + \tfrac{1}{2}\sin 2\delta\cos\epsilon\cos\sigma)g^2 - (\tfrac{1}{2}\sin 2\epsilon\sin\sigma)h^2 + (\cos 2\delta\cos\epsilon\cos\sigma + \tfrac{1}{2}\sin 2\delta\sin 2\epsilon\sin\sigma)fg + (\cos\delta\cos 2\epsilon\sin\sigma + \sin\delta\sin\epsilon\cos\sigma)fh + (\sin\delta\cos 2\epsilon\sin\sigma - \cos\delta\sin\epsilon\cos\sigma)gh\right]$$

$$G_{xz} = m\left[(\tfrac{1}{2}\cos^2\delta\sin 2\epsilon\cos\sigma + \tfrac{1}{2}\sin 2\delta\cos\epsilon\sin\sigma)f^2 + (\tfrac{1}{2}\sin^2\delta\sin 2\epsilon\cos\sigma - \tfrac{1}{2}\sin 2\delta\cos\epsilon\sin\sigma)g^2 - (\tfrac{1}{2}\sin 2\epsilon\cos\sigma)h^2 + (\tfrac{1}{2}\sin 2\delta\sin 2\epsilon\cos\sigma - \cos 2\delta\cos\epsilon\sin\sigma)fg + (\sin\delta\cos 2\epsilon\cos\sigma + \cos\delta\sin\epsilon\sin\sigma)gh + (\cos\delta\cos 2\epsilon\cos\sigma - \sin\delta\sin\epsilon\sin\sigma)fh\right]$$

$$G_{yy} = m\left\{\left[\cos^2\delta(1-\sin^2\epsilon\sin^2\sigma) + \sin^2\delta\sin^2\sigma + \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma\right]f^2 + \left[\sin^2\delta(1-\sin^2\epsilon\sin^2\sigma) + \cos^2\delta\sin^2\sigma - \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma\right]g^2 + (\sin^2\epsilon + \cos^2\epsilon\cos^2\sigma)h^2 + \left[\sin 2\delta(\cos^2\sigma - \sin^2\epsilon\sin^2\sigma) - \cos 2\delta\sin\epsilon\sin 2\sigma\right]fg + (\sin\delta\cos\epsilon\sin 2\sigma - \cos\delta\sin 2\epsilon\sin^2\sigma)fh + (\cos\delta\cos\epsilon\sin 2\sigma + \sin\delta\sin 2\epsilon\sin^2\sigma)gh\right\}$$

$$G_{yz} = m\{[\tfrac{1}{2}\sin 2\sigma(\cos^2\delta\sin^2\epsilon + \sin^2\delta) - \tfrac{1}{2}\sin 2\delta\sin\epsilon\cos 2\sigma]f^2$$
$$+ [\tfrac{1}{2}\sin 2\sigma(\sin^2\delta\sin^2\epsilon - \cos^2\delta) + \tfrac{1}{2}\sin 2\delta\sin\epsilon\cos 2\sigma]g^2$$
$$+ (\tfrac{1}{2}\cos^2\epsilon\sin 2\sigma)h^2 + (\tfrac{1}{2}\sin\delta\sin 2\epsilon\sin 2\sigma + \cos\delta\cos\epsilon\cos 2\sigma)gh$$
$$+ [\tfrac{1}{2}\sin 2\delta(\sin^2\epsilon + 1)\sin 2\sigma + \sin\epsilon(\cos 2\delta\cos^2\sigma + \cos^2\delta\cos 2\sigma)]fg$$
$$+ (\tfrac{1}{2}\cos\delta\sin 2\epsilon\sin 2\sigma - \sin\delta\cos\epsilon\cos 2\sigma)fh\}$$

$$G_{zz} = m\{[\cos^2\delta(1 - \sin^2\epsilon\cos^2\sigma) + \sin^2\delta\cos^2\sigma - \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma]f^2$$
$$+ [\sin^2\delta(1 - \sin^2\epsilon\cos^2\sigma) + \cos^2\delta\cos^2\sigma + \tfrac{1}{2}\sin 2\delta\sin\epsilon\sin 2\sigma]g^2$$
$$+ (\sin^2\epsilon + \cos^2\epsilon\sin^2\sigma)h^2 + (\cos\delta\cos\epsilon\sin 2\sigma - \sin\delta\sin 2\epsilon\cos^2\sigma)gh$$
$$+ [\sin 2\delta(1 - \cos^2\sigma[\sin^2\epsilon + 1]) + \cos 2\delta\sin\epsilon\sin 2\sigma]fg$$
$$- (\sin\delta\cos\epsilon\sin 2\sigma + \cos\delta\sin 2\epsilon\cos^2\sigma)fh\}$$

3. Typical Aerodynamic Coefficients

Given below are two typical aerodynamic coefficients which were derived according to the procedure outlined in Section IV. All other coefficients can be derived in the same way.

$$C_{y_{p_1}} = \cos\delta\cos\epsilon\, C_{y_{p_2}} + (\cos\delta\sin\epsilon\cos\sigma + \sin\delta\sin\sigma)C_{y_{r_2}}$$
$$+ \frac{2\lambda_3}{\ell}[g(\cos\epsilon\sin\sigma) - h(\sin\delta\sin\epsilon\sin\sigma + \cos\delta\cos\sigma)]C_{y_{\beta_2}} \qquad (A\text{-}3)$$

$$C_{m_{\dot\alpha_1}} = C_{m_{\dot\alpha_2}}\lambda_1(\cos\delta\cos\epsilon\cos^2\sigma + \tfrac{1}{4}\sin\delta\sin 2\epsilon\sin 2\sigma) + C_{x_{\dot\alpha_2}}(\tfrac{\lambda_2}{\ell})$$
$$[h(\cos\delta\cos^2\epsilon\cos\sigma) + f(\tfrac{1}{2}\sin 2\epsilon\cos\sigma)] + C_{z_{\dot\alpha_2}}(\tfrac{\lambda_2}{\ell})[\tfrac{h}{2}(\cos\delta$$
$$\sin 2\epsilon\cos^2\sigma + \sin\delta\cos\epsilon\sin 2\sigma) - f(\cos^2\epsilon\cos^2\sigma)] \qquad (A\text{-}4)$$

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

REDUCING THE VOLUME OF COMPUTATION IN MONTE CARLO MISSILE FLIGHT SIMULATION

Prepared by:	Carl G. Looney, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Mathematics Department University of Toledo

Assignment:

(Air Force Base)	Eglin Air Force Base
(Laboratory)	Armament Laboratory
(Division)	Analysis Division
(Branch)	Missiles and Guns Analysis Branch

USAF Research Colleague:	Aaron D. Brinson
Date:	19 August 1977
Contract No.:	F44620-75-C-0031

REDUCING THE VOLUME OF COMPUTATION IN MONTE CARLO MISSILE FLIGHT SIMULATION

by
Carl G. Looney

ABSTRACT

A type of air-to-air heat-seeking missile is being evaluated by the USAF by means of a large scale computer program in which the thrust, seeker action, guidance controls, and associated parameters are mathematically simulated. Since some input variables have noise (random) components and are actually random variables, the output values have distributions which are unknown, and thus a Monte Carlo method is used. The known input random variables are sampled and fed through the program sufficiently many times to get convergence of the probability-of-kill for a missile launch from a particular area of an elevation plane which contains the target.

Two approaches are identified in this report for reducing the volume of computation. One is that of using random differential equations. The other (more practical, straight-forward) is one of passing the means and means plus or minus one standard deviation through the computational system and then fitting cumulative Weibull distributions to outputs of interest. Then samples can be drawn cheaply from these output distributions.

ACKNOWLEDGMENT

The author feels appreciation and gratitude for the opportunity, accommodation, and help presented to him this summer by the following people and organizations:

Fred O'Brien and Auburn University (for a smooth administration of the program); the Air Force Office of Scientific Research and Col. David England (for presenting the opportunity); Aaron Brinson (my research colleague who explained the simulations to me and gave me many results of his experience with the missile data and the computer program): Lt. Greg Keethler (for explaining certain missile hardware and for accessing test data recorded on tape); Jeff Kirkland (for helping with my computer programming needs); to the people in the Missile and Guns Analysis Branch of the Analysis Division, Armament Laboratory, Eglin AFB (for providing me with an office and computational facilities); and to the Commanding Officers of Eglin AFB (for my accommodations in the Visiting Officer's Quarters) and Armament Development Test Center (for full cooperation with this summer program). I want to thank both ASEE for their part in the program and Mrs. Jeanne Ferguson for her expert typing of this report.

1. INTRODUCTION

The Analysis Division of the Armament Laboratory at Eglin AFB simulates the flights of air-to-air missiles launched from a fixed altitude toward a target (at the same altitude) which is undergoing a preset target maneuver. The missile thrust, seeker action, and guidance controls are simulated mathematically, and the positions and velocities of the missile and target, as well as the pitch and yaw of the missile are computed time-incrementwise until the time of closest approach of the missile to the target. The lastly computed values are then fed into a routine (ENDGAME) which computes the probability of a kill of the target.

Certain input variables contain noise and thus are actually random variables. Therefore, any deterministic (e.g., mean) values put into the computer program yield output values which are unsatisfactory in that sampled random inputs yield sample outputs which are distributed in some manner around the deterministic values. Monte Carlo methods are used, i.e., samples are drawn (generated via random number generating subroutines) from the known distributions of the input random values and the computed outputs give sample distributions of the final values. These final values are generated and fed into the ENDGAME routine until convergence of the p_k (probability of kill) for launches from the given position.

Many launches are required (usually more than 30) for each p_k. Dozens of p_k's are needed for each p_k-envelope at a given altitude (with fixed target power and maneuver), so that dozens of p_k-envelopes are needed. But each launch uses approximately a thousand or more time increments, and at each increment 50 differential equations must be integrated and program bookeeping operations updated. Thus these p_k-envelopes are very costly (one may require the solution of 30 million differential equations).

The problem of reducing the volume of computation has been considered by various persons and technical contractors, but to date no generally satisfactory methods have been found.

2. OBJECTIVES

The purpose of research on this problem is to identify any possibly fruitful methods for reducing the volume of computation required to graph p_K-envelopes. The simulation equations can not be changed, of course, so that any acceptable method for computational reduction must either circumvent the purely Monte Carlo method, use more efficient algorithms, or "tighten up" the program. The program is already fairly tight and efficient algorithms are used for solving the differential equations.

3. THE MATHEMATICAL PROCESS FLOW

The computational flow of the simulation process may be considered to be a black box of unknown operations into which are put deterministic parameters $p_1, \ldots, p_m$, random parameters $U_1, \ldots, U_q$, and initial conditions of the variables of interest, $x_1, \ldots, x_n$. The outputs are updated values of these inputs at the next time increment and are random variables $X_1, \ldots, X_n$ because the U_j's have entered the scheme.

Diagram 1 shows the flow. The initial inputs (p_i°'s, U_j°'s and x_k°'s) enter on the left-hand side of the box, which represents a large scale computational system of such complexity that it is not known in functional description. The updated outputs on the right-hand side are the X_k's (the kinematics of the missile) and any new random parameters $U_{q+1}, \ldots, U_{q+r}$.

The outputs of interest are functions

$$x_k^{i+1} = f_k\,(p_1^\circ, \ldots, p_m^\circ, U_1^\circ, \ldots, U_q^\circ; U_{q+1}, \ldots, U_{q+r}, x_1^i, \ldots, x_n^i),\ 1 \le k \le n.$$

To gain knowledge of the total function $F = f_1, \ldots, f_n)$ we must sample its behavior.

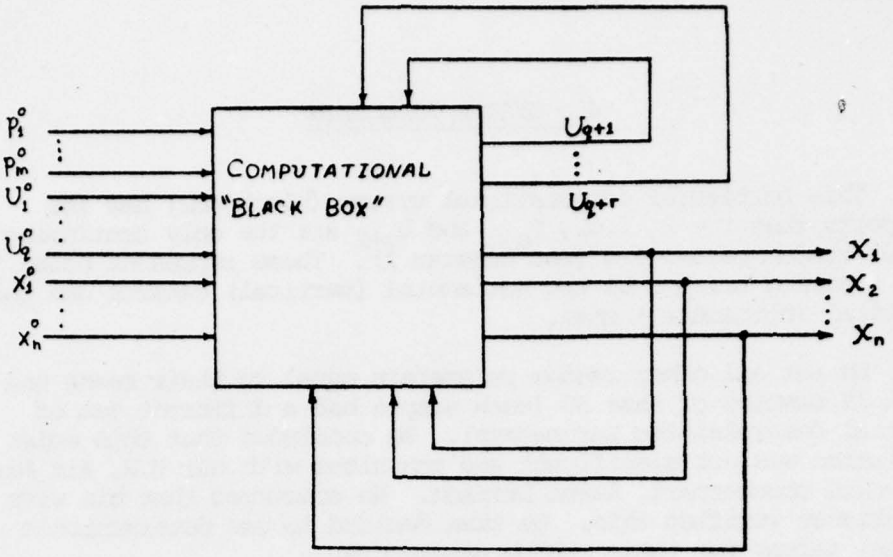


DIAGRAM 1: COMPUTATIONAL PROCESS FLOW

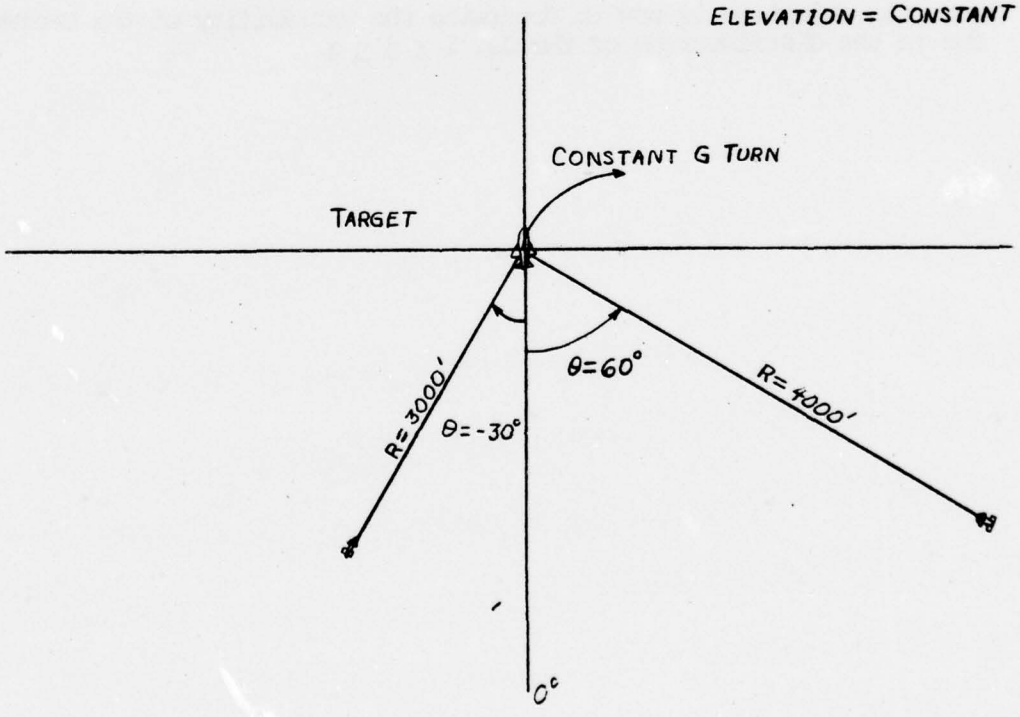


DIAGRAM 2: LAUNCH POINTS IN ELEVATION PLANE

4. SYSTEM REDUCTION

This particular computational system (black box) has the property that $r = 2$, i.e., U_{q+1} and U_{q+2} are the only continuing input random parameters (see Diagram I). These represent noise in the steering torques on the horizontal (vertical) canards due the vertical (horizontal) ones.

We set all other random parameters equal to their means and took 12 samples of size 30 (each sample had a different set of initial deterministic parameters). We concluded that this noise variation was not significant and consulted with our U.S. Air Force research counterpart, Aaron Brinson. He concurred that his wide experience verified this. We then decided to use deterministic (mean) values for these random parameters.

The system is now considered to be a vector function $\underline{x} = F \equiv (f_1, \ldots, f_n)$ with random initial conditions, which are on parameters only and not on the initial variables of interest.

Each set of values chosen for the random initial parameters fixes a computational system $F = (f_1, \ldots, f_n)$ through which we feed the variables of interest starting with the deterministic values $x_1, \ldots, x_n$ (once a fixed set of values $p_1, \ldots, p_m$ is given). We will call each such fixed system a "C-box".

Our objective is now to determine the variability of the C-boxes due to the distributions of the U_j, $1 \leq j \leq q$.

5. RANDOM AND DETERMINISTIC INITIAL PARAMETERS

The reduced system has 5 initial random parameters. All are Gaussian except U_5, which is uniform. Sampling and graphing of a few hundred histograms of sample sizes 50, 30 and 25 shows that the output X_k's are not Gaussian. Some are skewed to appear exponential, some are strongly skewed the other way, and some are in between (approximately Gaussian); but a few are bimodal (2 "humps").

$U_1 \equiv$ temperature-of-the-propellant is the source of greatest variability in the output X_k's according to sampling and reasoning.

The other random initial parameters can not be ranked in order of variability caused to the outputs, since the magnitudes of such effects depend on the conditions given by the p_1^o's.

$U_2 \equiv$ maximum-tracking-rate-achievable-by-the-missile causes greater output variance if the target is in afterburner power and is pulling a high g turn in the plane at shorter ranges and/or higher magnitudes for off-tail angle (see Diagram 2) θ.

$U_3 \equiv$ guidance-initiation-start-up-time (the amplifier warms up after launch time) contributes significantly to output variation when the overtake velocity is high at launch time. There is some significance of the noise at low altitudes (fast burn of the propellant gives high "instantaneous" thrust), short range launches ($\leq$ 1000 feet) and high magnitude of off-tail angle for high powered targets pulling high g turns.

$U_4 \equiv$ maximum-steering-torque-magnitude possible-on-canards causes significant output noise for high maneuvering target under high power for shorter ranges and/or high magnitude of off-tail angles θ.

$U_5 \equiv$ saturation-range (the range at which the seeker is close enough to become saturated and unable to follow target) yields significant noise only when target is under afterburner power. This one is uniformly, rather than Gaussianly, distributed.

The parameters of importance are p_1 = range, p_2 = off-tail-angle, and p_3 = altitude, all at launch time, along with p_4 = power-of-target (cruise, military, or afterburner) and p_5 = maneuver-of-target. The last 2 parameters are preset.

Once values are chosen for $p_1, \ldots, p_5$ and random values are generated for $U_1, \ldots, U_5$, a particular C-box $F = (f_1, \ldots, f_n)$ is fixed and is a deterministic vector function which takes the variables $x_1^o, \ldots, x_n^o$ into the final set of positions and velocities $x_1^f, \ldots, x_n^f$ at the discrete time increment of closest approach of the missile to the target. However, the random nature of the U's causes F to vary. It has been stated that the variations of these U's fix various F's that "skew-up" the outputs $X_1, \ldots, X_n$ for a fixed input $x_1^o, \ldots, x_n^o$.

6. COMPARISONS OF OUTPUT SAMPLES

The first two "samples" were taken by fixing the altitude at 10,000 feet, target power at afterburner, and target maneuver at a 5g right turn in the elevation plane. The range varied as 2000 to 6000 feet by increments of 1000 feet, while the off-tail angle varied as -90° to 90° by 15° increments. The first sample had all noises set to their mean values, while the second one had the maximum tracking rate U_2 set to the mean minus 2 standard deviations (over 98% of randomly generated values for this parameter would exceed this value).

Comparisons of output values showed that range (miss distance), missile pitch, yaw, and velocity magnitude (all at output time of closest approach) had no significant differences. The greatest differences were 1°, 1 foot, and 2 ft/sec. However, over 90% were the same when rounded to the nearest integer.

Two more "samples" were taken as above with only the altitude changed for both to be 40,000 feet. The results were the same except for the fact that about 4% of the miss distances had considerable differences. However, these were quite large miss distances so that a miss would be obtained in either case. In the other cases, the differences were negligible. The greater differences occurred at negative angles of magnitudes greater than or equal to 60°.

Then we put altitude = 10,000, range = 2000 to 4000 by 2000, off-tail angle = -75° to 75° by 30°, and fixed all noises at their means except the guidance start-up time U_3 (amplifier warm-up time) which varied according to its Gaussian distribution. A sample of 25 launches was made from each position with the result that the output data showed some small variation at certain launch points but virtually none at others. The greatest variation was essentially negligible at this altitude for the target pulling a 5g right turn under afterburner power. For shorter ranges, we would expect greater variation, and this should be sampled when time permits. We note that the overtake velocity was low in the above sample.

The next run was with altitude = 10,000 feet, target in afterburner power with no maneuvering, range = 2000, 4000 by 2000, and off-tail angle = -90°, 90° by 15°. All noises were put at their mean values except U_4, the maximum torque, which was set in 3 different runs at mean, mean minus 2 standard deviations, and mean plus 2 standard deviations, respectively. The miss distances and pitch and yaw of the missile were essentially the

same, with some error in the velocity of the missile at the final time increment. However, this error was negligible when considered as relative error (these velocities are in the neighborhood of 2000 ft/sec). Since 99% of all U_4 values would fall between the bounds given above, we conclude that this noise source is effectively insignificant at the ranges run. Very low or very high altitudes may affect the output and should be set in future sampling runs.

U_5 is the saturation range (at which the heat seeker can no longer function properly because of excessive signal) is a significant source of output noise. Here, we used test data already existing and consulted Aaron Brinson (who developed the simulations) to conclude this whenever the target is in afterburner power.

7. A METHOD FOR SHAPING THE OUTPUT DISTRIBUTIONS

The fault in the deterministic method can easily be seen by considering a standard normal variable Z which is put through the simple C-box $f(z) = z^2$. The mean of Z is 0 so that $f(0) = 0$ is the computed deterministic value when we use the mean as a deterministic representation of Z. However, values generated randomly from Z and fed through the C-box will have a chi-square distribution as limit with mean value of 1 and all outputs will be greater than 0. Thus the deterministic output value 0 would be unsatisfactory for use as a representative value for the outputs. What we need here, in addition to the value $0 = f(0)$, is an approximate shape of the frequency distribution of the outputs.

For the purpose of explanation, let us suppose that we have a situation where the noise in the output is practically all caused by, say, U_1 and U_5. First we set the other random initial parameters at their mean values. Let σ_1^2 and σ_5^2, μ_1 and μ_5 be their variances and means, respectively. Letting F designate the vector valued function for the C-box, we will now write $F(U_1, U_5)$ for the output vector to show the dependence on these random initial conditions with all other inputs being fixed.

In this particular case we compute the output vectors $\underline{x}^1 = F(\mu_1, \mu_5)$, $\underline{x}^2 = F(\mu_1 + \sigma_1, \mu_5 + \sigma_5)$, $\underline{x}^3 = F(\mu_1 + \sigma_1, \mu_5 - \sigma_5)$, $\underline{x}^4 = F(\mu_1 - \sigma_1, \mu_5 + \sigma_5)$, and $\underline{x}^5 = F(\mu_1 - \sigma_1, \mu_5 - \sigma_5)$. From the 5 sets of output data and the knowledge of the proportional frequencies that it represents, we want to construct the approximate shape of the output distributions for $X_1, \ldots, X_5$.

We need to use a cumulative distribution function of wide applicability which can fit distributions whose probability density functions are skewed drastically either to left or right or are approximately symmetrical. The Weibull [2] distribution is best suited to this need (we ignore the fact that finite probability exists for arbitrarily large numbers since, as in the exponential and Gaussian distributions, these finite probabilities are approximately 0 except on a closed bounded interval).

Thus we approximate X_i $(1 \leq i \leq 5)$ with a random variable W_i (a,b,c), which has cumulative distribution function $1 - \exp[-a(x-c)^b]$ and density function $ab(x-c)^{b-1} \exp[-a(x-c)^b]$, $x \geq c$; $a,b > 0$. We observe that for $b = 1$ this gives an exponential density with positive values on $x \geq c$.

We recall a theorem from mathematical statistics [4] which states that any probability distribution which is nonzero on a closed bounded interval is determined uniquely by all its moments. Since our computerized distributions have values which are nonzero only on closed bounded intervals, the theorem applies to our approximate distributions. We note that the

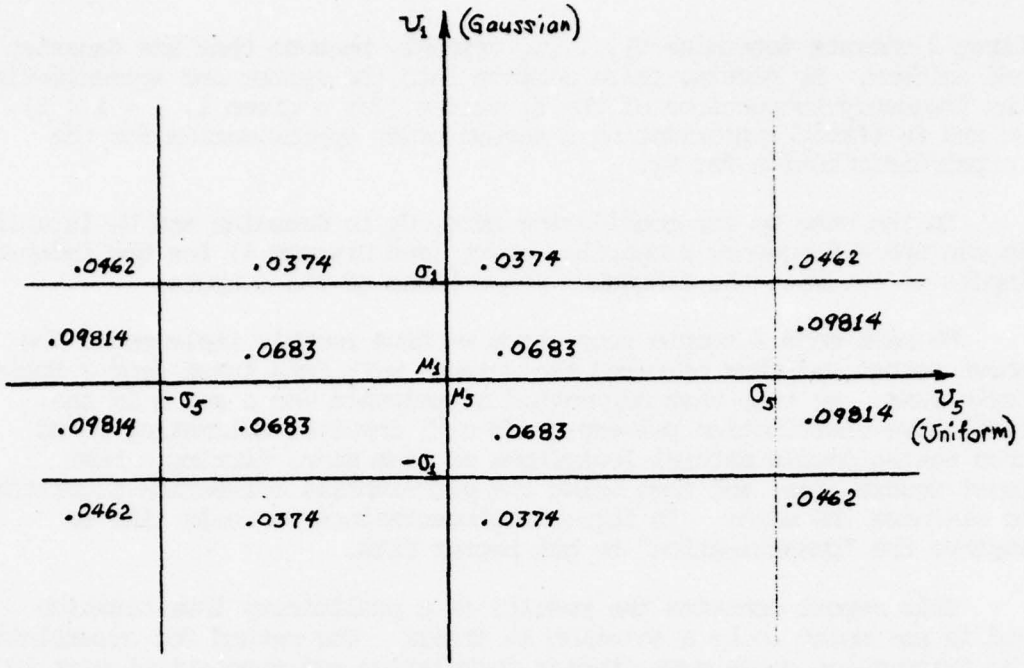


DIAGRAM 3: FREQUENCY PROPORTIONAL CONTRIBUTIONS

first 2 moments determine $U_1,\ldots,U_5$ uniquely because they are Gaussian and uniform. By putting these moments into the system and approximating the frequency proportions of the X_i values (for a given i, $1 < i < 5$), we are in effect approximating a second order approximation for the output distribution for X_i.

In the case we are considering here, U_1 is Gaussian and U_5 is uniform. We can use a frequency proportion chart (see Diagram 3) for the independent inputs to estimate the frequency proportions of the outputs.

We have made 2 sample runs where we have roughly implemented the above method and then compared the outputs with data taken from 2 Monte Carlo runs. We note that our method of solution for a and b in the cumulative distribution $y=1-\exp[-a(x-c)^b]$ involves estimating c and then taking double natural logarithms of each side, fitting a best least square line, and then using the differential correction algorithm to minimize the error. In future implementations we would like to improve the "guesstimation" to get better fits.

This report contains the results of a preliminary investigation and is not meant to be a solution in itself. Our method for accumulating the frequencies needs more precise formulation and theoretical work in which statistical bounds can be given. Our fitting procedure needs improvement. Even so, the results of the runs, with R and θ chosen randomly, show that this approach is superior to that of computing with means only. It is not known how precise the Monte Carlo method is for samples of this size, since the inputs drawn at random do not approximate Gaussian and uniform distributions too well (n should be considerably larger).

The data used to compute the means, standard deviations, and $F(\mu_1, \mu_2)$ have been rounded off somewhat.

The data denoted by "xxx" was not obtained since our method for computing a and b did not converge for our estimate of c.

We need to stress the fact that these runs constitute a very rough implementation of our method and are not a fair test. However, they demonstrate the possibilities.

COMPARISON 1, R = 4000 ft., 0 = 60°

METHOD DATA		MISS DISTANCE	PITCH	YAW	VELOCITY
μ	(n=30)	6.99	- .47	78.4	1833.3
σ		4.9	3.0	6.96	55.5
MONTE CARLO DATA					
μ	(n=24)	6.42	-1.2	75.3	1804.6
σ		2.9	.38	1.2	15.3
F (μ_1, μ_5)		4.	-1.	75.	1800.

COMPARISON 2, R = 2000 ft., 0 = -15°

METHOD DATA					
μ	(n=30)	48.4	.9	xxxx	2309.2
σ		27.7	7.9	xxxx	115.5
MONTE CARLO DATA					
μ	(n=30)	60.9	-4.1	3.6	2234.9
σ		32.0	2.2	3.4	24.6
F (μ_1, μ_5)		51.	-4.	4.	2242.

For n sources of noise which cause significant output variation, this method requires $1 + 2^n$ launches (if triple combinations of $\pm \sigma_i$'s were used, $1 + 3^n$ launches would be needed, etc.). For small $n \leq 4$, this method seems superior to Monte Carlo in efficiency.

8. AN ALTERNATE METHOD

Let $\underline{x} = (x_1, \ldots, x_n)$ depend on t and $\underline{F} = (f_1, \ldots, f_n)$. In terms of the differential equations of the system, we have [1], [3]:

$$\dot{\underline{X}}(t) = \underline{F}(\underline{X}(t), t) \text{ for } t \text{ in } [0, T] \text{ and } \underline{F}(\underline{X}(0), 0) = \underline{F}^{\circ}.$$

This is a nonlinear random differential equation with initial random function conditions (a generalization of random linear differential equations with random coefficients [4]).

A vector $\underline{x}^{\circ}$ which is passed consecutively through such a system until the time increments total up to T, where at each pass the solution yields a random variable, will yield an output random variable solution which can then be sampled. However, each incremental solution requires a number of iterations, which one would hope would take a lesser volume of computation required than the taking of the required number of launches for the Monte Carlo method.

There is currently a gap between the theory (existence and uniqueness) of solutions and the computation of approximative solutions. In general, there is no numerical method yet devised, and the special known cases do not seem to apply here.

An incremental solution involves the solution of a stochastic integral:

$$\underline{X}(t+h) = \underline{X}(t) + \int_t^{t+h} \underline{F}(\underline{X}(t), t)\, dt.$$

9. NOTES, CONCLUSIONS, AND RECOMMENDATIONS

We do not expect a viable method of solution for the random differential equation to be available in the near future. The method of fixing C-boxes at plus and minus standard deviations holds much promise for the short and intermediate range future (at least). We recommend further research into the method and techniques as well as implementation wherever the inefficiency of Monte Carlo methods is to be traded off against a (possibly) slight degradation in accuracy (and wherever the deterministic method is not satisfactory).

In the current simulation computer program we recommend:

i) the use of a subroutine to use the time and date (from the computer) to seed the random number generating subroutines, since our samples showed that the same set of random numbers is being used on each run,

ii) a scheme to increase the time increment size where guidance control, GC, is at a low level, since our test runs have indicated that time can be saved without loss of too much accuracy by a multistage decision loop for changing the time increment to one of several values (Aaron Brinson identified GC as the significant determiner of time increment size for us), since a saving of 1 second per launch would yield a 7.1% savings,

iii) that analytic solutions to differential equations be used where possible (suggested by Lt Greg Keethler),

iv) that the variances of the outputs of the C-box method be diminished by truncating the fitted Weibull distributions so that they will have positive weighting on closed bounded intervals of shorter length, in order to model the real situation more accurately,

v) research a method for fitting the Weibull distributions by solving for the parameter c as well as for a and b, since our present "guesstimation" of c is a strong source of error in the shape of the output distributions,

vi) investigate methods for $n>3$ to accumulate the frequencies of the outputs around the $1 + 2^n$ computed points (see the last paragraph of Section 7),

vii) develop a computer subroutine to be used in conjunction with Monte Carlo simulations which will implement the results of parts iv, v, and vi above.

REFERENCES

1. Astrom, Karl J., Introduction to Stochastic Control Theory, Academic Press, New York, 1970.

2. Johnson, Norman L. and Leone, Fred C., Statistics and Experimental Design in Engineering and the Physical Sciences, Vol. I, Wiley & Sons, New York, 1964.

3. Soong, T.T., Random Differential Equations in Science and Engineering, Academic Press, New York, 1973.

4. Wilks, Samuel S., Mathematical Statistics, Wiley & Sons, New York, 1962.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

FLUSH MOUNTED AIRCRAFT ANTENNAS FOR SATELLITE COMMUNICATION; STUDY OF MULTIPATH PERFORMANCE

Prepared by:	Vaughn P. Cable, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Electrical and Computer Engineering California State University Northridge
Assignment:	
(Air Force Base) (Laboratory) (Division) (Branch)	Hanscom AFB Rome Air Development Laboratory Electromagnetic Sciences Division Antennas and RF Components Branch
USAF Research Colleague	Robert J. Mailloux, Ph.D.
Date:	August 19, 1977
Contract No.:	F44620-75-C-0031

FLUSH MOUNTED AIRCRAFT ANTENNAS FOR SATELLITE COMMUNICATION. A STUDY OF MULTIPATH PERFORMANCE

by

Vaughn P. Cable, Ph.D.

ABSTRACT

This report presents the results of a preliminary investigation on the performance of airborne phased arrays in an aircraft/satellite multipath environment. A two-dimensional geometry is assumed and a model is developed which predicts the multipath performance of flush (conformal) arrays mounted on the upper part of the main aircraft fuselage. The diffracted multipath energy is reflected off the earth's surface and compared to the energy in the direct path as a function of satellite position (zenith to horizon). These first results indicate improved performance (reduced level of multipath) for larger arrays, but that improvement is also a function of placement of the array. When a specific requirement for "rejecting" multipath is set, say 20 dB down, the optimum array position is on the side of the aircraft (45^{o} off zenith). Here, the conventional phased array also has its best scan capability (zenith to horizon). However, if a 40 dB threshold for rejecting multipath is needed, the optimum position appears to be at the top (zenith). The conventional array in this case has a somewhat limited scan coverage, but a hybrid (slow wave mode) is mentioned as a possible solution to expanding this coverage to the horizon.

Acknowledgements

The author is sincerely greatful to the USAF and the ASEE for their support of the Summer Faculty Research Program, for without it, this effort would not have been possible. The author would also like to thank Dr. Allan Schell for his encouragement and especially Dr. Robert Mailloux, Dr. Ronald Fante, Col William Goggins, Dr. Robert Papa and Dr. Richard Mack for their helpful and illuminating discussions. Last, but certainly not least, the author would like to thank Raymond Cormier and Alice Cahill for their willing assistance and especially Louise Tanguay for her typing of this report.

TABLE OF CONTENTS

I. Introduction

The purpose of this investigation is to do an initial study of the multipath problem between aircraft and satellite and as a result develop a multipath model to measure performance of flush mounted (conformal) aircraft antennas in a multipath environment. Aperture size and placement of these antennas are to be parameters in this model and the model will then be tested on various antenna configurations in order to develop this tool for optimizing such antennas for superior multipath performance.

The overall multipath problem deals with the many simultaneous paths which a signal can travel between sender and receiver. These paths for the aircraft/satellite link can be quite complicated, especially when wings, nose and tail section scattering and diffraction are taken into account. The major multipath system is however, between these fuselage points, down to the surface of the earth (e.g. ocean) and back up to the satellite. This latter multipath mechanism will be studied in detail in this report for a somewhat simplified aircraft geometry.

This report will only be considering the performance of flush mounted aperture antennas. The extreme importance of these antennas is to be found in their smooth aerodynamic characteristics. They are also capable of either electronic or mechanical (or both) scanning. Single arrays or combinations of 2 or more such arrays are expected to give hemispherical coverage from zenith to the horizon above the aircraft.

Considerable literature is available on both phased arrays [1,2,3] and satellite coverage antennas [4] and also on multipath effects on LOS and ionospheric communication as well as for radar applications [5,6,7,8]. This report however, seeks to combine portions of these disciplines to develop a tool which would aid the

design of specialized airborne phased arrays with optimum multipath rejection in a generalized multipath environment.

The antenna system and multipath geometries and the development of the model are described in detail in Chapter II. Computed results based on this model are presented and discussed in Chapter III. Finally, the important results are summarized and recommendations for further development of the model are called for in the last Chapter (IV). Suggestions are also made in this last chapter for new possible methods for improving the multipath performance of phase scanned arrays on cylinders.

II. DEVELOPMENT OF THE MODEL

The antenna configuration modeled in this study is a 2-dimensional set of closely spaced apertures (e.g., open ended waveguides) flush mounted on an aircraft fuselage. This conformal array is to be electronically scanned for near-hemispherical coverage. A similar array has been investigated by Mailloux and Mavroides [9].

The basic approach used in this investigation considers a simplified (2-dimensional) antenna-fuselage geometry with beam scanning only in the circumferential ϕ direction. Figure 1 illustrates this 2-dimensional geometry. This view shows an

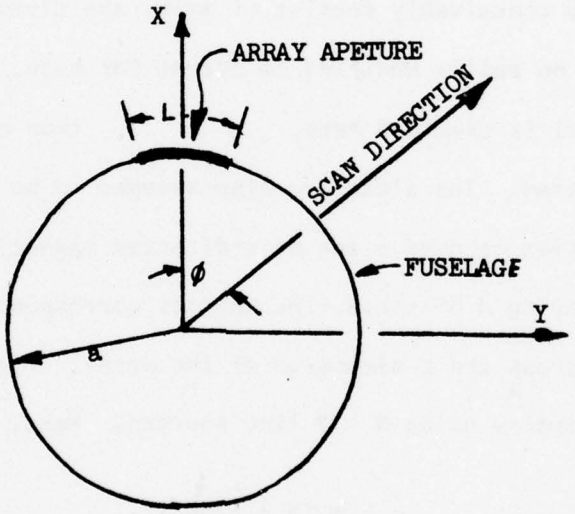


Figure 1. Cross section view of flush mounted array on aircraft fuselage ($\phi_s = 0^o$)

array with circumferential length L wavelengths positioned on "top" of the fuselage. Arbitrary circumferential position is specified in terms of the

array angular bisector ϕ_s for $0^\circ \leq \phi_s \leq 45^\circ$. Two array "widths" are investigated in this report; L = 3.12 wavelengths and L = 7.8 wavelengths. The fuselage is assumed to be a perfectly conducting right circular cylinder of radius "a" wavelengths and infinitely long in the $\pm$ z directions. Wings, nose and tail sections are not considered however, more complex geometries which include these characteristics can be treated with these same techniques

Application of the equivalence principle [10] allows the array apertures, or "slots," to be replaced by a magnetic source distribution $\vec{M}$ on the fuselage surface. This equivalent source distribution is given by

$$\vec{M}(Q') = \vec{E}(Q') \times \hat{n}' \quad , \tag{1}$$

where Q' is a source point and $\hat{n}'$ is the unit outward normal vector at Q'. The array could conceivably consist of axial and circumferential slots and this model can be easily modified to account for both. Only the axial slot case (TE_z excitation) is examined here, thus only ϕ directed polarization need be considered. The slots are also assumed to be "narrow" or thin hence, the $\vec{M}(Q')$ distribution becomes a set of z directed magnetic line sources. The number N and spacing d of these line sources correspond to the number and spacing of the slots across the L dimension of the array. Figure 2 shows the 2-dimensional geometry using N = 7 line sources. Here, L is given by

$$L = (N - 1)\, b/\lambda \quad , \tag{2}$$

where b is the narrow dimension of the waveguide apertures in meters and λ is the operating wavelength in meters.

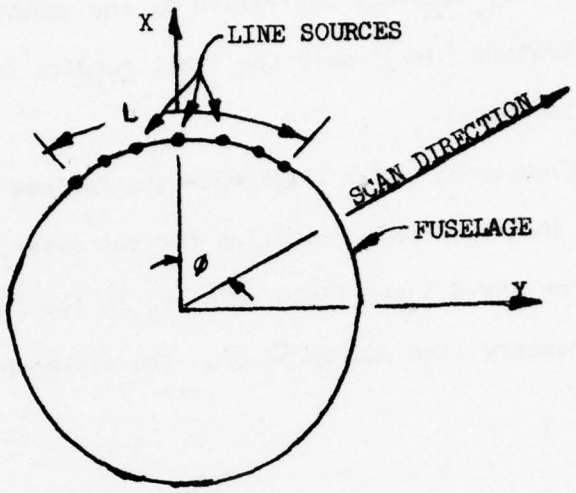


Figure 2. Magnetic line source model for N thin axial slots across array aperture (N = 7). (Note, $\phi_s = 0^\circ$).

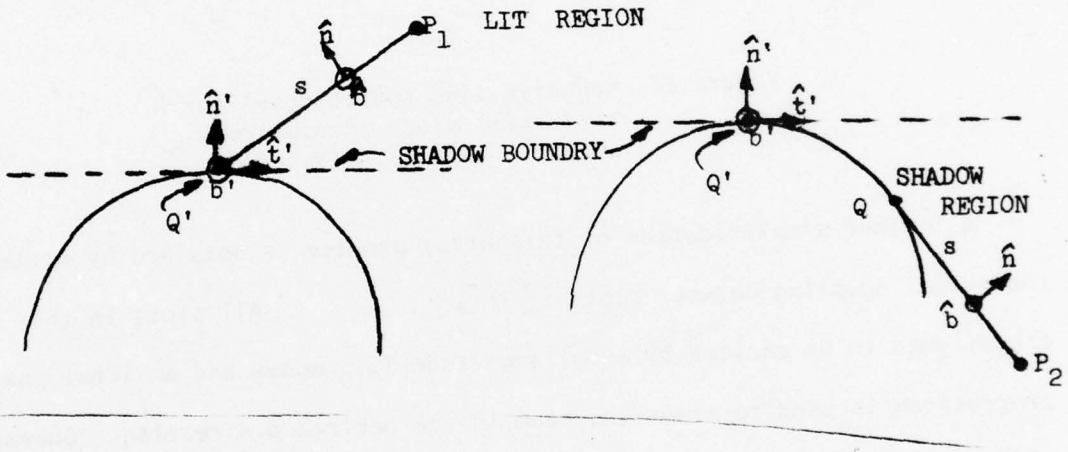
A further simplification of this array problem is obtained by neglecting the mutual coupling between slots [11]. All slots in this case are assumed to be excited by equal amplitude TE_{10} modes and an ideal phase progression is used to steer the beam in the desired ø direction. Curvature of the array is also accounted for in scanning the array.

The field pattern of a slotted infinite circular cylinder can be solved by the classical eigenfunction method [12]. However, the Geometric Theory of Diffraction [13] (GTD) approach is applied here in order to handle more complex shapes if need be (e.g. wings, engines, nose & fail section). The GTD solutions used in this development follows the latest work of Pathak and Kouyoumjian [14, 15].

A. GTD Formulation of the Fields.

A more complete derivation of the general equations can be found in the literature [16]; only the final results for the 2-dimensional case are presented here.

Figures 3a and b illustrate the "primed" source coordinate system as well as the field point position for two cases; a.) the field point P_1 in lit region and b.) the field point P_2 in the shadow region, both for a single elementary line source at Q'. The distance S in Fig. 3a.



a. P_1 in lit region.

b. P_2 in shadow region.

Figure 3. Localized surface coordinate system at source point Q'. Note, $\hat{n}'$, $\hat{t}'$ and $\hat{n}'$ are in the plane of the figure and $\hat{b}' = \hat{t}' \times \hat{n}'$. Also, $\hat{b} = \hat{b}'$ (non-torsional case).

corresponds to the direct ray path length from Q' to P_1, whereas, S in Fig. 3b is the ray path length from the tangent point Q (where the diffracted ray is "shed" toward P_2) to P_2.

The incremental electric field intensity $d\vec{E}(P)$ at any point P due to and incremental source $d\vec{M}(Q')$ at Q' is given by

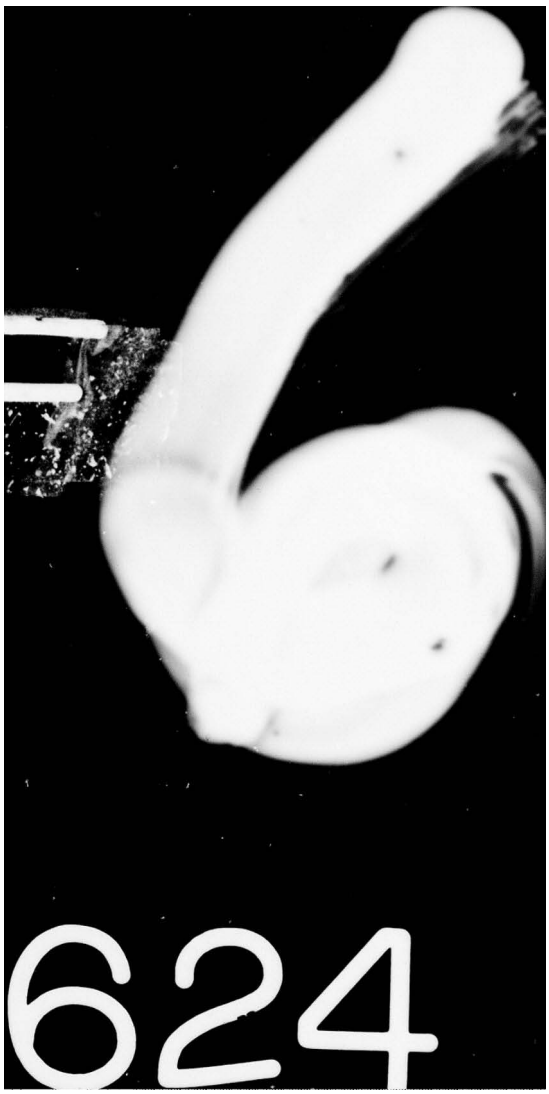
$$d\vec{E}(P) = -k\sqrt{\frac{\mu}{8\pi k}}\, d\vec{M}(Q') \cdot \left[\hat{b}'\hat{n}\, F(Q;Q) + \hat{t}'\cdot\hat{b}\, G(Q;Q)\right] \frac{e^{-jks}}{\sqrt{s}} \qquad (3)$$

where $k = \omega\sqrt{\mu\epsilon}$ radians/meter. The F and G functions appearing in Eq. 3 are diffraction coefficients and explicit expressions for these functions can be found in the reference [17]. Note, that when the direct ray field, P_1 is in the lit region, Q = Q'. An exception to this occurs if the diffracted rays completely encircle a closed convex surface. The geometries considered here are of sufficient size ($a = 25\lambda$) to neglect these encircling rays and Figure 4 illustrates the ray fields to be considered in this study, i.e., the direct ray field $\vec{E}_d$ and the two diffracted fields $\vec{E}_1$ (CW diffracted field) and $\vec{E}_2$ (CCW diffracted field). It is assumed that the $\vec{E}_1$ and $\vec{E}_2$ ray fields account for all "multipath" energy.

Recall, the assumed array consists of thin axial slots, hence, Eq. 3 need not be integrated. The field is that of a single line source distribution which is uniform in the z direction and the total radiated field of the array is the algebraic summation of the $\vec{E}_d$, $\vec{E}_1$ and $\vec{E}_2$ fields from each individual line source.

AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

4 OF 6
AD
A051 624


624

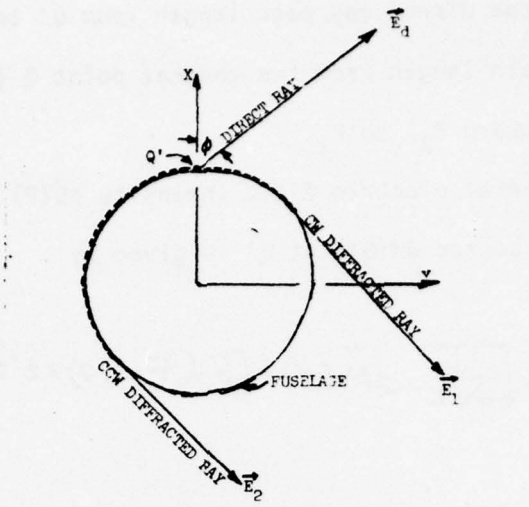


Figure 4. Significant ray fields used in the multipath problem.

B. The Multipath Geometry

Figure 5 illustrates arbitrary aircraft and satellite positions over a spherical reflecting earth.

Calculation of the direct ray path electric field intensity $\vec{E}_d$ is a straight forward application of Eq. 3 with $S = S_d$. The multipath calculations, however, are complicated by the reflection process at the earth's surface. These reflections are dependent on a number of factors including the nature of the surface, the incidence or grazing angle (Ψ) and the polarization of the incoming rays.

Random or rough surface scattering [18] will certainly modify the amplitudes and phases of these multipath rays. however, these factors will only reduce the "expected" amount of energy in the multipath rays system. Only the case of specular scattering (reflection) from a smooth lossy dielectric surface will be considered in this report. This is, in a certain sense, a worst case analysis of this multipath problem. The incorporation of a rough

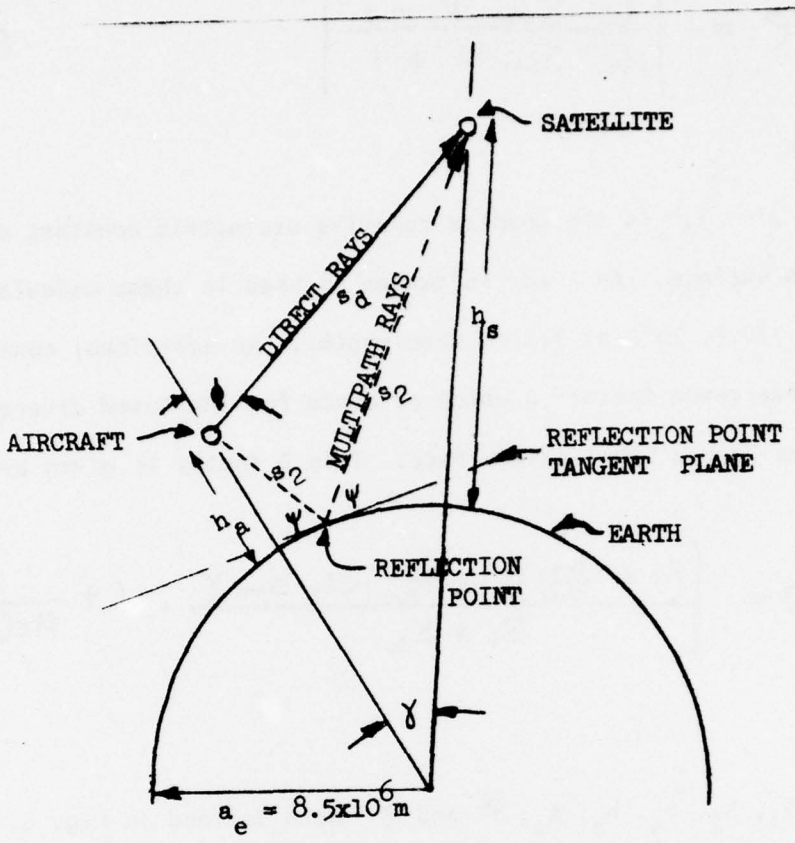


Figure 5. Multipath geometry for aircraft satellite communication path. Note, effective radius of earth a_e accounts for atmospheric refraction effects where $a_e = 4/3$ x (actual radius of earth).

surface reflection model would be straight forward based on the appropriate model appearing in the literature [19].

The reflection coefficient used in this study is derived from the Fresnel equation [20] for a smooth planar surface. A simplified expression for the magnitude is given here for the required parallel polarization case only,

$$\rho = \left| \frac{\sqrt{\epsilon}\,\sin\Psi - 1}{\sqrt{\epsilon}\,\sin\Psi + 1} \right| \qquad (4)$$

where $\epsilon = \epsilon' - j\epsilon''$ is the complex relative dielectric constant of the reflection surface. An ϵ for saltwater is used in these calculations ($\epsilon = 65 - j30.7$, $28^{\circ}C$ at 3.2 cm wavelength). An additional consideration is the "divergence factor" D which accounts for increased divergence of reflections from a spherical surface. This D factor is given by

$$D = \left[\frac{\left(1 + \frac{h_s}{a_e}\right)\left(1 + \frac{h_a}{a_e}\right) a_e \sin\gamma}{S_1 + S_2} \cdot \frac{1 + 2S_1S_2}{a_e(S_1+S_2)\sin\Psi} \right]^{-\frac{1}{2}} \qquad (5)$$

where

S_1, S_2, h_s, h_a, A_e, Ψ and γ were defined in Fig. 5. Equations for calculating the reflection point and thus the grazing angle Ψ for the above expressions are given in the references [21,22].

The magnitudes of the two multipath electric field intensities $|\vec{E}_{1S}|$ and $|\vec{E}_{2S}|$ arriving at the satellite are calculated as follows;

$$|\vec{E}_{1S}| = \rho D |\vec{E}_1| \qquad (6)$$

and

$$|\vec{E}_{2S}| = \rho D |\vec{E}_2| \qquad (7)$$

where $S = S_1 + S_2$ in Eq 3 .

A measure of the "interference" at the satellite is difficult to define without knowledge of the type of communication being attempted (e.g., digital, voice, non-coherent, coherent, etc.). A convenient measure for the general

case might best given by relating the power densities for the two arriving fields. The multipath rejection ratio (MRR) defined by

$$MRR = -10 \log \frac{|E_{1s}|^2 + |E_{2s}|^2}{|E_d|^2} \qquad (8)$$

provides this relationship. Note, $\vec{E}_d$, $\vec{E}_{1S}$ and $\vec{E}_{2S}$ in Eq 8 are the electric field intensities at the satellite. A copy of the computer program which calculates Eqs 3 and 8 verses scan angle ϕ for $0^\circ \leq \phi \leq 90^\circ$ (zenith to horizon) appears in Appendix A and results are given in the next section for various antenna parameters (ϕ_S, L) and aircraft-satellite altitudes (h_S, h_a).

III. Results

Initial calculations which verify beam shapes and steerability are shown in Figures 7-12 for two cases, $L = 3.12\lambda$ (9 slots) and $L = 7.8\lambda$ (21 slots). The arrays were positioned on top of the cylinder ($\phi_s = 0^o$) for these pattern calculations and the main beam was placed in a sequence of directions ($0^o \leq \phi \leq 90^o$). Figures 6-9 are E-plane power patterns for $L = 3.12\lambda$ with $\phi = 0^o$, 30^o, 60^o and 90^o and Figs 10-12 are E-plane power patterns for $L = 7.8\lambda$ with $\phi = 0^o$, 45^o and 90^o.

The "broadside" main beam ($\phi = 0^o$) cases shown in Figs 7 and 10 illustrate the typical main beam and near sidelobe patterns of a uniformly illuminated aperture with half-power beam widths of 15^o for $L = 3.12\lambda$ and 7^o for $L = 7.8\lambda$. The sidelobes nearing the "horizon" however, merge together and form rather broad backlobes beneath the aircraft (cylinder). This is characteristic of all scan angles shown and appears to be the major contribution from the fuselage. This region of the radiation pattern is of concern since these are the angles which contribute directly to the multipath ray system. Recall here also that these relative patterns are insensitive to cylinder diameter for this polarization whereas, the TM_z case would not be. The cylinder radius used for all calculations in this report is $a = 25\lambda$. This same contribution from the fuselage begins to "contaminate" the main beam for the lower scan angles ($\phi > 45^o$). The main beam is in fact, near destroyed for the $\phi \geq 60$ case for the smaller array (Fig 8). It is clear from these patterns that this type of phased array cannot be successfully scanned near "endfire." One obvious reason is that the apertureson the far side are shadowed by the cylinder. Figures 9 and 12 show that the actual peak of the beam is not at 90^o for $\phi = 90^o$; calculated peaks actually occurred at 74^o for $L = 3.12\lambda$ and 79^o for $L = 7.8$. The peak intensities of the main beams

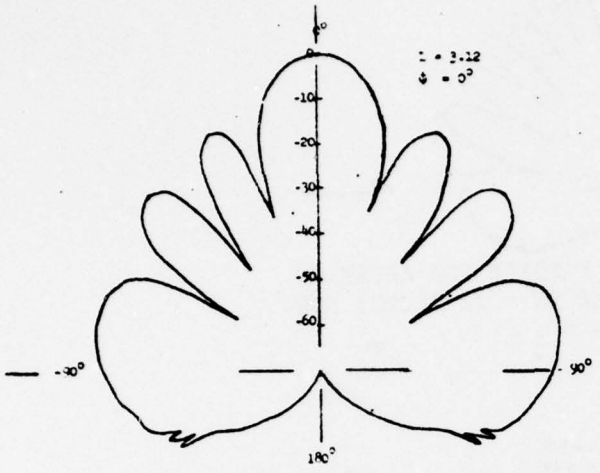


FIGURE 6.

E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 9 SLOTS) FOR SCAN ANGLE $\phi = 0^{\circ}$.

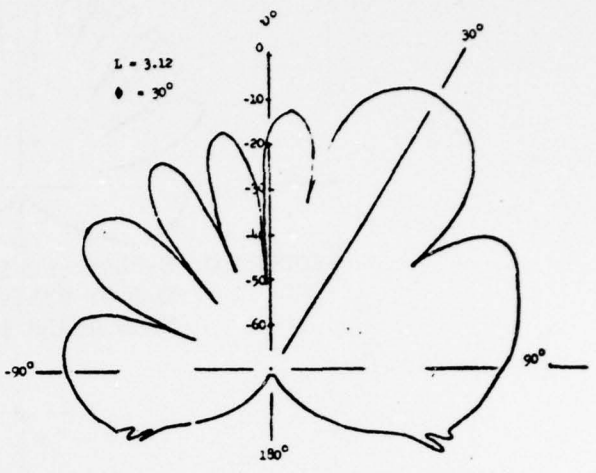


FIGURE 7.

E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 9 SLOTS) FOR SCAN ANGLE $\phi = 30^{\circ}$.

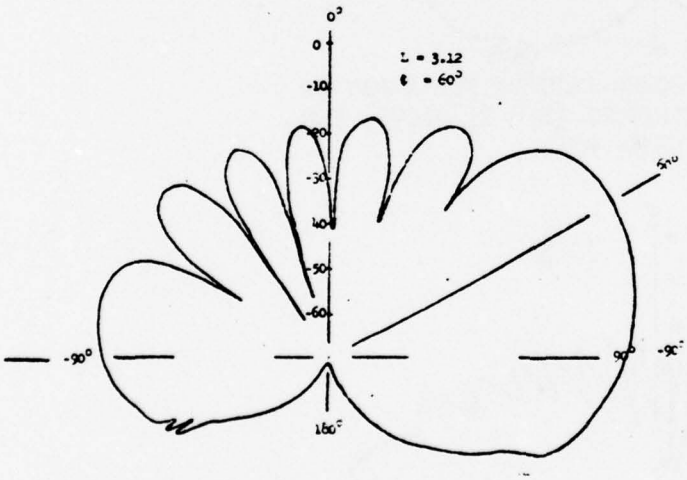


FIGURE 8.

E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N=9 SLOTS) FOR SCAN ANGLE $\phi = 60^{\circ}$.

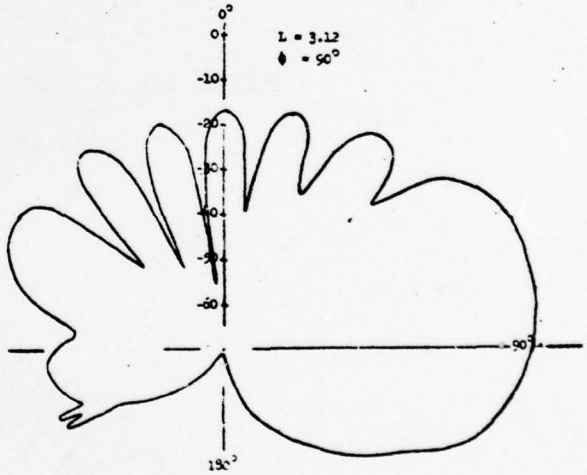


FIGURE 9.

E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 9 SLOTS) FOR SCAN ANGLE $\phi = 90^{\circ}$.

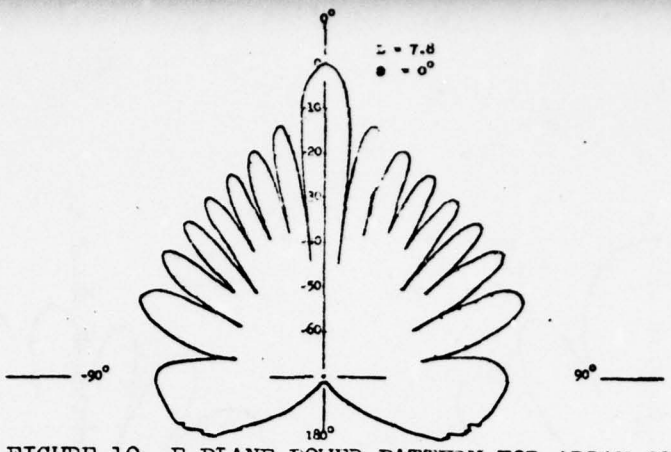


FIGURE 10. E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 21 SLOTS) FOR SCAN ANGLE $\phi = 0^{\circ}$.

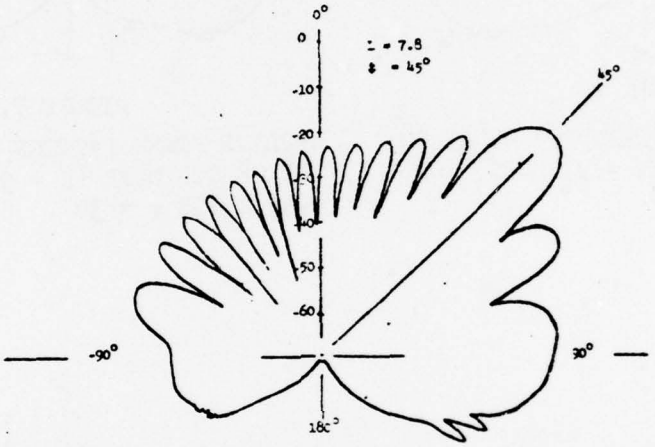


FIGURE 11. E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 21 SLOTS) FOR SCAN ANGLE $\phi = 45^{\circ}$.

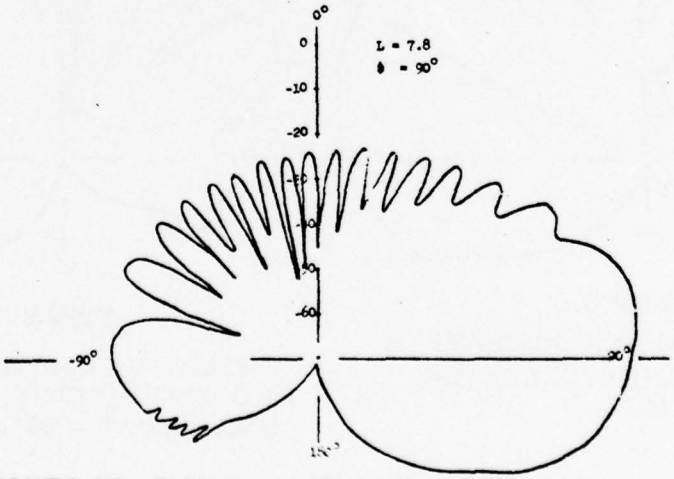


FIGURE 12. E-PLANE POWER PATTERN FOR ARRAY ON TOP OF CYLINDER (N = 21 SLOTS) FOR SCAN ANGLE $\phi = 90^{\circ}$.

in these two cases were only .66 dB ($L = 3.12\lambda$) and 1.42 dB ($L = 7.8\lambda$) below peak intensity for broadside scan ($\phi = 0^{o}$) and the intensities at 90^{o} (the horizon) were 3.12 dB ($L = 3.12\lambda$) and 3.26 dB ($L = 7.8\lambda$) below peak at broadside. Notice that this power reduction at the horizon seems to agree with the shadowing effect mentioned earlier. This endifre condition and a possible solution to the problem will be discussed in some detail in the conclusions of this report.

Figures 13 and 14 show calculated multipath rejection ratios (MRR) verse array scan angle (ϕ) for various aircraft/satellite altitudes for two arrays ($L = 3.12\lambda$ and $L = 7.8\lambda$) and two array positions ($\phi_s = 0^{o}$ and $\phi_s = 45^{o}$). Also, the aircraft is assumed to be in level flight for all calculations throughout this report. Figure 13 ($L = 3.12\lambda$, $\phi_s = 0^{o}$) shows the relatively small improvement in MRR (2-3 dB) which is gained by increasing the aircraft altitude from .1 km (317 ft) to 10 km (31,700 ft), in this case for a fixed satellite altitude of 200 km (120 mi). This same figure also shows the significant improvement obtained when the satellite is in a synchronous orbit, i.e., 36000 km (22,500 mi). Here, the aircraft altitude is held at 10 km. Note, the greatest improvement in this case (aircraft alt = 10 km) occurs at the horizon (50 dB). Figure 14 shows the variations in MRR for the larger array ($L = 7.8\lambda$) with the array located at $\phi_s = 45^{o}$. The satellite altitudes used here are 200 km and 36000 km and the aircraft altitude is fixed at .1 km. The improvement is again greatest at the horizon (40 dB) for the higher satellite.

The basic conclusion to be drawn from Figs 13 and 14 is that an increase in altitude for either end of the communication link can only improve the overall MRR. The remaining calculations will therefore show only data for the worst case condition of a low flying aircraft ($h_a = .1$ km) and a low fast moving satellite ($h_s = 200$ km).

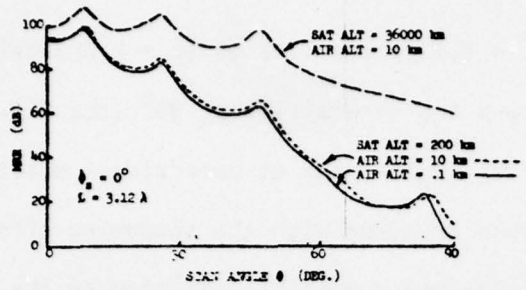


FIGURE 13. MRR vs ϕ FOR 1) h_s = 36000 km, h_a = 10 km, 2) h_s = 200 km, h_a = 10km, 3) h_s = 200 km, h_a = .1 km (ϕ_s = 0°).

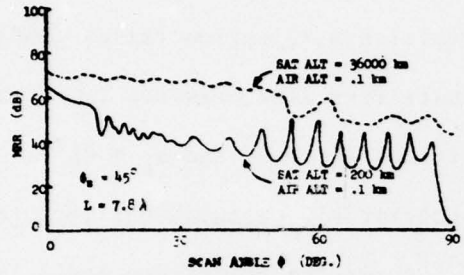


FIGURE 14. MRR vs ϕ FOR 1) h_s = 36000 km, h_a = .1 km, 2) h_s = 200 km, h_a = .1 km (ϕ_s = 45°).

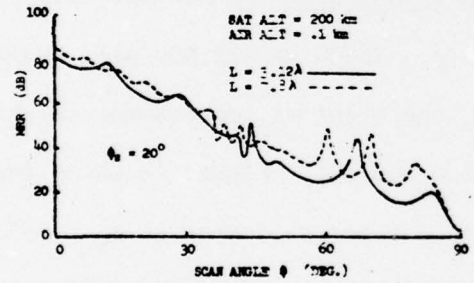


FIGURE 17. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 20°).

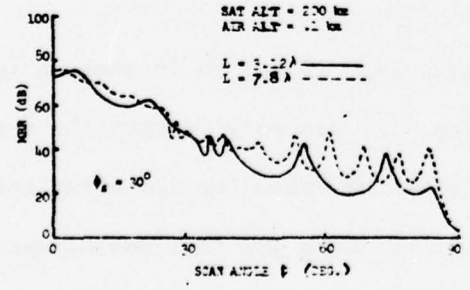


FIGURE 18. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 30°).

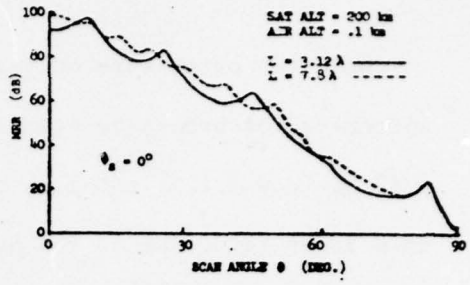


FIGURE 15. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 0°).

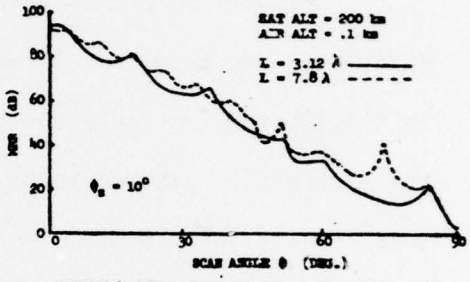


FIGURE 16. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 10°).

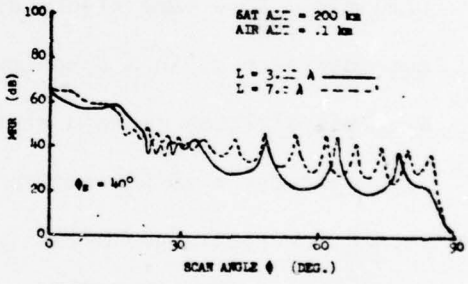


FIGURE 19. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 40°).

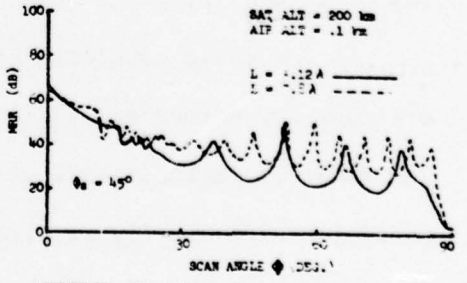


FIGURE 20. MRR vs ϕ FOR 1) L = 3.12λ, 2) L = 7.8λ (ϕ_s = 45°).

Figures 15-20 display curves of MRR vs ϕ for $\phi_s = 0^o$, 10^o, 20^o, 30^o, 40^o, and 45^o; each figure is plotted for two array sizes, $L = 3.12\lambda$ and $L = 7.8\lambda$. One of the major characteristics indicated by these curves is that the larger array has a higher (better) overall MRR. This is of course expected since the larger array has approximately 3 dB more directivity (normalized to a common H-plane beamwidth). Notice here that for $\phi_s = 0^o$, the improvement is in fact close to 3 dB over most of the scan range. This general improvement in MRR with the larger array is even more pronounced as the arrays are moved around to the side; e.g., for $\phi_s = 45^o$ (Fig 20) an increase of 10 dB is seen over significant sectors of ϕ. There are exceptions to this trend and these occur in relatively narrow isolated portions of the ϕ range. The MRR exhibits a series of peaks in these narrow ranges for both array sizes. These "periodic" peaks become more prominent as $\phi_s \rightarrow 45^o$ and are without a doubt related to aperture size similar to that of the radiation pattern sidelobe structure. The one rather broad peak consistently appearing near $\phi = 84^o$ in most of these figures is primarily due to Brewster angle phenomena and not sidelobe structure.

The other major trend showing up in the figures (Figs 15-20) is the marked decrease in MRR with ϕ. This however, is not unexpected since diffracted energy which contributes to the multipath system necessarily increases as ϕ is brought down to the horizon. Note, the greatest fall off of MRR(> 90 dB) occurs for $\phi_s = 0^o$ (Fig 15).

Let us define for convenience, an arbitrary MRR threshold for a reliable communication channel (i.e., signal to noise ratio). Consider two cases; MRR $\geq$ 40 dB for acceptable communication and MRR $\geq$ 20 dB for acceptable communication. These thresholds can now be used to define a <u>useable scan range</u> ϕ_{sr} for each antenna configuration and this ϕ_{sr} range is summarized here for comparison in Fig 21.

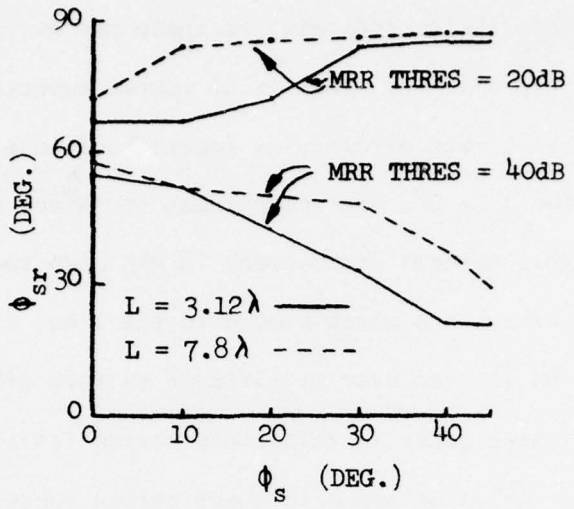


Figure 21. Useable scan range ϕ_{sr} vs ϕ_s for two MRR thresholds, 20 dB and 40 dB.

These results indicate that ϕ_{sr} decreases with increasing ϕ_s for MRR $\geq$ 40 dB but, increases for increasing ϕ_s for the less severe MRR $\geq$ 20 dB. This figure thus implies that for a communication channel which can withstand only 20 dB suppression of the multipath energy, the overall useable scan range for this aircraft system is greatest when the array is mounted on the side, say $\phi_s \geq 30^o$. There are many other considerations to be considered in designing such a system, but, given the constraints outlined here, data such as that shown in Fig 21 could be useful in optimizing placement of such an antenna.

The MRR behavior shown in the previous figures is obviously quite sensitive to the particular side and back lobe structure which directly affects the multipath system. Recall from Fig 20 ($\phi_s = 45^o$), that the MRR rapidly falls below 40 dB but the "average" remains > 20 dB over a greater ϕ_{sr} range than for the $\phi_s = 0^o$ case (Fig 15). Now, consider these same two cases at only the $\phi = 45^o$ scan angle.

The MRR for these cases differs significantly and even though the $\phi_s = 45^\circ$ case is scanned exactly broadside (its "best" scan angle), the MRR is 25-30 dB below that for the $\phi_s = 0^\circ$ case. It would appear for this case that the main beam directivity, in terms of beamwidth, is not as important in improving the MRR as the form and position of the sidelobe structure relative to scan angle.

IV. Summary and Recommendations

One of the more obvious trends which appears in these results is the effect which altitude has on multipath performance. This data shows that if aircraft altitude or satellite altitude (or both) is increased, the MRR improves (increases) in a generally monotonic fashion. This is due to the increased spreading of the multipath rays due to the divergence factor D. These results also indicate that an increase in the size of the array also improves the MRR. This may be somewhat misleading since this measure does not account for the difficulty in scanning larger conventional arrays, especially near endfire. More will be said about this later.

The results of the previous chapter also indicate an increased useable scan range (87^{o}) for side mounted arrays if the acceptable MRR threshold is set at 20 dB. However, for the 40 dB threshold, the data shows just the opposite to be true, i.e., the array in this case has its greatest scan range (58^{o}) when mounted directly on top of the fuselage. A possible solution for adequately covering the endifre scan range for this 40 dB case might be a hybrid structure which switches to a slow wave endfire mode for scan angles $\geq 60^{o}$. This array has been investigated by Mailloux [23] however, before a valid comparison can be made with data presented here, this slow wave excitation case needs to be evaluated using these same criteria. It should be mentioned that even this slow wave structure is apt to produce significant energy in the diffracted field. This modified approach may not then be the ultimate solution to the problem in terms of multipath performance.

Recall again that if the less severe threshold of 20 dB were acceptable, a large conventional side mounted ($\phi_s = 45^{o}$) array would be nearer the optimum

solution. It is possible however, that in certain cases 20 dB will not be sufficient rejection of the multipath energy. Other array aperture distributions might be adapted to produce minimal side lobes which contribute less to the multipath system (e.g., triangular or binomial distributions) but again, these non-uniform aperture distributions need to be evaluated under these same conditions before conclusions can be drawn.

A new possible solution to reducing this backlobe diffracted energy might be to use a "stop band" surface adjacent or near the aperture. This complex surface could be used to "detach" the diffracted energy, scattering it away from the fuselage, or even to absorb and dissipate the energy on the fuselage. Again, this scheme would need to be subjected to these same multipath performance criteria before a final "optimum" is found. As a final comment, it is entirely possible that the cylinder contribution to this deffracted field is so overwhelming that very few if any of these modifications would lead to significant improvements in multipath performance. This remains to be determined.

This study has succeeded in introducing a new tool for evaluating the multipath performance of flush mounted on-aircraft antennas and even with the simplified model used here, this approach provides at least a "first look" at the aircraft/satellite multipath problem in terms of parameters familiar to the antenna designer.

V. References

[1] N. Amitay, V. Galindo and C.P. Wu, Theory and Analysis of Phased Arrays, Wiley-Interscience, New York, 1972.

[2] A. A. Oliner and G. H. Knittel, Phased Array Antennas, Artech House, 1972.

[3] B. L. Diamond, "A Generalized Approach to Analysis of Infinite Planar Arrays," Proc. IEEE, Vol 56, Nov 1966, pp 1837-1851.

[4] J. P. Grabowski and F. L. Lanphear, "SHF High power Airborne Communications Antenna," Symposium on USAF Antenna R&D, Univ of Ill, Oct 1970.

[5] R. L. Fante, P.R. Franchi and R. L. Taylor, "Effect of Multipath on Height-Finding Radar Systems," In-House Report, RADC-TR-76-169, May 1976.

[6] M. W. Long, Radar Reflectivity of Land and Sea, Lexington Books, Lexington, MA 1975.

[7] G.T. Ruck, Radar Cross Section Handbook, Vol 2, Plenum Press, New York, 1970.

[8] K. Davies, Ionospheric Radio Propagation, NBS Monograph 80, April 1, 1965.

[9] R. J. Mailloux and W. G. Mavroides, "An Array Technique for Zenith to Horizon Coverage," In-house Report, RADC-TR-76-360, Nov 1976.

[10] R. F. Harington, Time-Harmonic Electromagnetic Fields, McGraw-Hill Co., New York, 1961.

[11] W. Wasylkiwskyj, "Mutual Coupling Effects on Semi-Infinite Arrays," IEEE Trans on AP, May 1973.

[12] R.F. Harington —

[13] J. B. Keller, "Geometric Theory of Diffraction," J. Opt. Soc Amer, Vol 52, No. 2, Feb 1962.

[14] P. H. Pathak and R.G. Kouyoumjian, "An Analysis of the Radiation from Apertures in Curved Surfaces by the Geometric Theory of Diffraction," Proc. IEEE, Vol 62, Nall, Nov 1974.

[15] P. H. Pathak, private communication, June 1977.

[16] P.H. Pathak and R. G. Kouyoumjian, —

[17] R. G. Kouyoumjian, "Introduction to the Geometric Theory of Diffraction," Short Course on GTD and Numerical Techniques in the Analysis of EM and Acoustic Radiation and Scattering, The Ohio State Univ., Sep 1974.

[18] M. W. Long, —

[19] G. T. Ruck, —

[20] D.E. Kerr, Propagation of Short Radio Waves, Dover Publications, New York, 1965.

[21] D. E. Kerr, —

[22] M. W. Long —

[23] R. J. Mailloux, "Phased Array Aircraft Antennas for Satellite Communication," to be published in Microwave Journal, Fall 1977.

APPENDIX A.

COMPUTER PROGRAM

Antenna Patterns

```
      PROGRAM GTOEL (INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)                    000100
      COMPLEX ETH,EPH,EDT,CJ                                                   000110
      DIMENSION ED(360),APHI(360)                                              000120
      COMMON /PLKABF/A,R,VSR,PHSR,ZS,JANT,SGN                                  000130
      COMMON /BLKBET/BETA                                                      000140
      COMMON /LONG/PATFAC                                                      000150
      CJ=(0.,1.)                                                               000160
      PI=3.14159265                                                            000170
      R0=5378E3*4/3                                                            000180
      READ(5,500)A,ARYAG,N,C,PHO                                               000190
  500 FORMAT(2F3.0,1I3,2F3.1)                                                  000200
      NP=N/2+1                                                                 000210
      JANT=1                                                                   000220
      BETA=0                                                                   000230
      ZS=0                                                                     000240
      THR=PI/2                                                                 000250
      B=A                                                                      000260
      AR=A*.0375                                                               000270
      ARYAGR=ARYAG*PI/180                                                      000280
      PHOR=PHO*PI/180                                                          000290
      ABR=PHOR-ARYAGR                                                          000300
      SPS=SIN(ARYAGR)                                                          000310
      CPS=COS(ARYAGR)                                                          000320
      DPHSR=D/A                                                                000330
      WRITE(6,502)A,ARYAG,N,C,PHO                                              000340
  502 FORMAT(1H1,"A="1F3.0,1X,"ARYPOS="1F3.0                                   000350
     2,1X,"N="1I2,1X,"SPACING="1F4.2,1X,"PHO="1F3.0)                           000360
      WRITE(6,504)                                                             000370
  504 FORMAT(1HB,3X,"PHI",7X,"EFIELD",4X,"PWRPAT")                             000380
      PHI=-180                                                                 000390
      EMAX=0                                                                   000400
      DO  900 IP=1,360                                                         000410
      PHI=PHI+1                                                                000420
      IF(ABS(PHI).LT..001)PHI=.001                                             000430
      PHR=PHI*PI/180                                                           000440
      SPH=SIN(PHR)                                                             000450
      CPH=COS(PHR)                                                             000460
      APHI(IP)=PHI                                                             000470
      IF(PHO.LT.0)ABR=(PHR-ARYAGR)                                             000480
      EDT=(0.,0.)                                                              000490
      PHSR=-NP*DPHSR+ARYAGR                                                    000500
      DO 880 IR=1,N                                                            000510
      NN=-NP+IR                                                                000520
      CORD=1.414*A*SQRT(1-COS(NN*DPHSR))                                       000530
      ARG=ABR-NN*DPHSR/2                                                       000540
      IF(NN.GE.0)ARG=-ARG                                                      000550
      DEL=2*PI*CORD*SIN(ARG)                                                   000560
      PHSR=PHSR+DPHSR                                                          000570
      VSR=PHSR                                                                 000580
      SGN=1                                                                    000590
      TEST=PHR-PHSR                                                            000600
      IF(TEST.LT.0)SGN=-1                                                      000610
      IF(ABS(TEST).LT.1E-5)SGN=1                                               000620
      CALL FUSLAG(ETH,EPH,THR,PHP)                                             000630
      EDT=EDT+EPH*CEXP(CJ*DEL)                                                 000640
  880 CONTINUE                                                                 000650
      E=CABS(EDT)                                                              000660
      ED(IP)=E*F                                                               000670
      EMAX=AMAX1(EMAX,ED(IP))                                                  000680
  900 CONTINUE                                                                 000690
      DO 920 IZ=1,360                                                          000700
      PD=10*ALOG10(ED(IZ)/EMAX)                                                000710
      WRITE(6,600)APHI(IZ),ED(IZ),PD                                           000720
  600 FORMAT(1X,1F7.2,1X,1E12.5,1X,1F7.2)                                      000730
  920 CONTINUE                                                                 000740
 1000 CONTINUE                                                                 000750
      STOP                                                                     000760
      END                                                                      000770
```

MRR Program

```
      PROGRAM STOOL (INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
      COMPLEX ETH,ERH,EDT,EP1T,EP2T,CJ,CE
      DIMENSION EP(91),EP1(91),EP2(91),APHI(91),SD(91),S1(91),S2(91)
      DIMENSION SI(91)
      COMMON /BLKARF/A,R,VSQ,FWSP,ZS,JANT,SGN
      COMMON /BLKBET/BETA
      COMMON /LONG/PATFAC
      CJ=(0.,1.)
      PI=3.14159265
      RO=6378E3*4/3
      READ(5,550)N9
  550 FORMAT(1I2)
      DO 1000 IF=1,N9
      READ(5,500)ALT,SALT,A,ARYAG,N,D,PHD,E1,E2
  500 FORMAT(2E7.0,2F3.0,1I3,2F3.1,2F5.1)
      ADR=SIN((PHD-ARYAG)*PI/180)
      SARO=SALT+RO
      ARO=ALT+RO
      NP=N/2+1
      CE=CSQRT(E1-CJ*E2)
      STR3=SQRT(2*RO*SALT)
      STRA=SQRT(2*RO*ALT)
      T=STRT(ALT/SALT)
      GT=ACOS(RO/(RO+ALT))
      RL=GT*RO
      JANT=1
      BETA=0
      ZS=0
      THR=PI/2
      B=A
      AR=4*.0375
      ARYAGR=ARYAG*PI/180
      SPS=SIN(ARYAGR)
      CPS=COS(ARYAGR)
      DPHSR=D/A
      WRITE(6,602)ALT,SALT,A,ARYAG,N,D,PHD
  602 FORMAT(1H1,"ALT="1E9.3,1X,"SALT="1E9.3,1X,"A="1F3.0,1X,"ARYAG="1F3.0,1X,
     2,1X,"N="1I2,1X,"SPACING="1F4.2,1X,"PHD="1F3.0)
      WRITE(6,640)E1,E2
  640 FORMAT(1X,"REL DIELECTRIC CONS="1F5.1,"-J"1F5.1)
      WRITE(6,604)
  604 FORMAT(1H0,3X,"PHI",4X,"PSI2",4X,"MRR",7X,"PD",7X,"EP1",6X,"EP2")
      PHI=-1
      EMAX=0
      DO 800 IP=1,91
      PHI=PHI+1
      IF(PHI.LT..0011)PHI=.001
      PHR=PHI*PI/180
      SPH=SIN(PHR)
      CPH=COS(PHR)
      APHI(IP)=PHI
      GAMA=PHR-ASIN(ARO*SPH/SARO)
      SD(IP)=(SALT+RO)/SPH*SIN(GAMA)
      R=GAMA*RO
      P=2/1.7320508*SQRT(RO*(ALT+SALT)+R/2*R/2)
      PSI=ACOS(2*RO*(SALT-ALT)+R/(P*P*P))
      R1=R/2+P*COS((PSI+PI)/3)
      R2=R-R1
      ZETA=R2/RO
      S1(IP)=SQRT(ARO*ARO+RO*RO-2*ARO*RO*COS(ZETA))
      S2(IP)=SQRT(SARO*SARO+RO*RO-2*SARO*RO*COS(GAMA-ZETA))
      PHRI=PI-ACOS((S1(IP)*S1(IP)+ARO*ARO-R1*R1)/(2*S1(IP)*ARO))
      SP1=R1/STRS
      SP2=R2/STRA
      PSI2=ATAN2((ALT+SALT)*((1-SP1*SP1)+T*T*(1-SP2*SP2)),R*(1+T*T))
      SI(IP)=PSI2*180/PI
      SP2=SIN(PSI2)
      CP2=COS(PSI2)
      RV=CABS((CE*SP2+1)/(CE*SP2+1))
      S12=S1(IP)+S2(IP)
      DP=AR*RO*S12*S12*CP2*SP2/((S1(IP)*(S2(IP)+RO*SP2)+S2(IP)*
     2(S1(IP)+RO*SP2))*ARO*SARO*SIN(GAMA))
      IF(PHD.LT.0)PHR=(PHR-ARYAGR)
      EDT=(0.,0.)
      EP1T=(0.,0.)
      EP2T=(0.,0.)
      PHSR=-NP*DPHSR+ARYAGR
      DO 880 IR=1,N
      NN=-N+IR
      CORD=1.414*A*SQRT(1-COS(NN*DPHSR))
      ARG=ADR*NN*DPHSR/2
      IF(NN.GE.0)ARG=-ARG
      DEL=2*PI*CORD*SIN(ARG)
      PHSR=PHSR+DPHSR
      VSR=PHSR
      SGN=1
```

```
      TEST=PHR-PHSR                                                     000960
      IF(TEST.LT.0)SGN=-1                                               000970
      IF(ABS(TEST).LT.1E-5)SGN=1                                        000980
      CALL FUSLAG(ETH,EPH,THR,PHP)                                      000990
      EDT=EDT+EPH*CEXP(CJ*DEL)                                          001000
      IF(R2.GE.RL)GO TO 880                                             001010
      SGN=1                                                             001020
      CALL FUSLAG(ETH,EPH,THR,PHP1)                                     001030
      EP1T=EP1T+EPH*CEXP(CJ*DEL)                                        001040
      SGN=-1                                                            001050
      CALL FUSLAG(ETH,EPH,THR,PHP1)                                     001060
      EP2T=EP2T+EPH*CEXP(CJ*DEL)                                        001070
  889 CONTINUE                                                          001080
      E=CABS(EDT)                                                       001090
      ED(IP)=E**2                                                       001100
      EV=CABS(EP1T)*RV                                                  001110
      EP1(IP)=EV*EV*DP                                                  001120
      EV=CABS(EP2T)*RV                                                  001130
      EP2(IP)=EV*EV*DP                                                  001140
  900 CONTINUE                                                          001150
      DO 920 IZ=1,91                                                    001160
      PO=ED(IZ)/SO(IZ)                                                  001170
      PR1=EP1(IZ)/(S1(IZ)+S2(IZ))                                       001180
      PR2=EP2(IZ)/(S1(IZ)+S2(IZ))                                       001190
      PHRR=10*ALOG10((PR1+PR2)/PO)                                      001200
      WRITE(6,600)APHI(IZ),SI(IZ),PHRR,ED(IZ),EP1(IZ),EP2(IZ)          001210
  600 FORMAT(1X,1PE7.2,1X,1PE7.2,1X,1PE7.2,1X,1PE7.2,1X,5E9.3)          001220
  920 CONTINUE                                                          001230
 1000 CONTINUE                                                          001240
      STOP                                                              001250
      END                                                               001260

      SUBROUTINE FUSLAG(ETH,EPH,THR,PHR)
      COMPLEX ETH,EPH,WOPE,EASE,YYZ,G,GB,EX,SFASE,FZ,FTAU,GTAU,GSKWTG,GJ
      COMMON /BLKBAS/A,B,VSP,PHSP,ZS,JANT,SGN
      COMMON /BLKSET/BETA
      COMMON /LONG/PATFAC
      DATA PI,TPI,PO2/3.14159265,6.2831853,57.29577958/
      DATA HPI/1.5707963257
      BLEN=.25
      PATFAC=1
      JANT=1
      BETAR=BETA/PDR
      CST=COS(THR)
      SNT=SIN(THR)
      SNP=SIN(PHR)
      CSP=COS(PHR)
      CJ=(0.,1.)
      EX=CEXP(CJ*TPI*ZS*CST)
      SML=.0001
      VR=ATAN2(A*SNP,B*CSP)
      IF(VSP*SGN.GT.SML.AND.SGN*VR.LT.-SML)VR=VR+SGN*TPI
      IF(SGN*(VR-VSP).LT.SGN*(-SML))GO TO 1
      IF(SGN*(VR-VSP)*DEG.GT.115) GO TO 1
      SNPSP=SIN(PHSP)
      CSPSP=COS(PHSP)
      SNVS=SIN(VSR)
      CSVS=COS(VSR)
      CSRT=COS(BETAR)
      SNT=SIN(BETAR)
      XS=A*CSVS
      YS=B*SNVS
      FASE=CEXP(CJ*TPI*(XS*SNT*CSP+YS*SNT*SNP))
      SFASE=FASE*EX
      XX=A*SNVS
      YY=B*CSVS
      CSPREL=B*CSVS*CSP+A*SNVS*SNP
      CSPREL=ABS(CSPREL)/SQRT(XX*XX+YY*YY)
      VRE=VR-SGN*HPI
      RG=(XX*XX+YY*YY)**1.5
      RETAU=RG/A/B
      GTAU=RETAU/SNT/SNT
      FACKRG=(1/RT/RG)**(2./3)
      FACKRG=FACKRG/SNT/SNT
      SNVE=SIN(VRE)
      CSVE=COS(VRE)
      XDE=A*SNVE
      YDE=B*CSVE
      FACTOR=1/(PI*RG)**(1./2)
      RDG=(XDE*YDE+YDE*YDE)**1.5
      EGS=GS/A/B/SNT/SNT
      GS=HP
      IF(SGN*(VRE-VRA).LT.0) GO TO 111
      CALL SWAP(SKWTG,THR,VRP,VRS)
      SKWTG=-SKWTG
      FZL=1
      FZLG=CGT(1-CSPREL*CSPREL)
      GTLG=1
      IF(ABS(SKWTG).GT.1.E-5)GTLG=SNT*CSPREL/FACTOR/ABS(SKWTG)
```

```
      CALL ETNT(ANS,VSR,VSR)
      DL=ANS/ABS(SNT)
      ZT=DL*CST
      GO TO 102
  101 CALL EKARG(SKWIG,THR,VSR,VFR)
      CALL ETNT(ANS,VSR,VFR)
      EZLC=(360/PI)**(1./6)
      ETAC=EZLC
      GTLC=EZLC
      DL=ANS/SNT
      ZT=DL*CST
  102 CONTINUE
      XT=AR*CSVR
      YT=AR*SNVR
      EASE=CEXP(CJ*TPI*(YT*SNT*CSP+XT*SNT*SNP+ZT*CST))
      XYZ=EASE*CEXP(-CJ*TPI*FL)
      GSKWIG=CONJG(G(SKWIG))
      EZ=XYZ*GSKWIG*EZLC/2
      ETAU=XYZ*GSKWIG*ETLC/2
      GTAU=2C*XYZ*(-CJ)/2*FACTOR*CONJG(G3(SKWIG))*GTLC/PGTAU
      IF(SKWIG.GE.0) GO TO 202
      ETH=SNT*SFASE*CABS(SNT*GTAU)
      EPH=CSPT*SNT*SFASE*CABS(EZ)-SNBT*CST*SGN*SFASE*(CABS(ETAU)+CABS(FA
     ZCKRR*ETAU/2))
      RETURN
  202 CONTINUE
      HOPE=-CJ*.5*CEXP(CJ*PI/6)
      ETH=SNBT*SNT*GTAU
      EPH=-SGN*SNBT*CST*ETAU+CSBT*SNT*EZ-SGN*SNBT*CST*HOPE*FACKRR*PTAU
      RETURN
    1 CONTINUE
      CALL SGTP(VSR,THR,PHR,ETH,EPH,BETA2)
      ETH=ETH*EX
      EPH=EPH*EX
      RETURN
      END

      SUBROUTINE SGTP(VSR,THR,PHR,ETH,EPH,BETAR)
      COMPLEX EZH,ETAUS,ETAUH,FCORR,CJ,F,HOPE,ALH(2,2),DET(2,2),AL(2,2)
      COMPLEX ETH,EPH
      COMMON /BLKARR/A,B,VSCR,PHSR,ZS,JANT,SGN
      COMMON /GTP/AS,MODE,ID,IDTX
      COMMON /LONG/EATFAC
      DIMENSION CPAR(2)
      DATA CPAR/1.01879,3.2482/
      DATA PI,TPI,DPR/3.14159265,6.2831853,57.29577958/
      DATA HPI/1.570796325/
      IDTX=1
      MODE=2
      CJ=(0.0,1.0)
      AS=SGN*90
      SNT=SIN(THR)
      CST=COS(THR)
      SNP=SIN(PHR)
      CSP=COS(PHR)
      VR=ATAN2(A*SNP,B*CSP)
      IF(VR.LT.0)VR=VR+TPI
      IF(SGN.GT.0)VFR=VR-HPI
      IF(SGN.LT.0)VFR=VR-3*HPI
      VSS=VSR
      IF(SGN*VFR.LT.SGN*VSS)VFR=VFR+SGN*TPI
      IF(SGN.LT.0)IDTX=2
      VSSD=VSS*DPR
      VFRD=VFR*DPR
      TANA=TAN(AS)
      SNA=SIN(AS)
      CSA=COS(AS)
      SNVS=SIN(VSS)
      CSVS=COS(VSS)
      SNV=SIN(VFR)
      CSV=COS(VFR)
      XD=AR*CSV
      YD=AR*SNV
      ID=3
      CALL QINT(VSS,VFR,VT)
      SS=ABS(VT/SNA)
      ZD=SS*CSA
      SNAB=SIN(AS-HPI)
      CONEA=A*A*SNVS*SNVS+B*B*CSVS*CSVS
      DTB=CON**.5
      BC=DTB/A/B/SNA/SNA
      FACMEG=1/(PI*BC)**(2./3)
      FACKBC=FACMEG/SNT/SNT
      BT=DTB/A/B/SNAB/SNAB
      IF(ABS(AS-HPI).LE.1E-4)BT=BC*1E15
      IF(ABS(AS-HPI*0I).LE.1E-4)BT=BC*1E15
      CALL LAUNCH(ALH,BC,BT)
```

```
      CON=A**3*SNV*SNV+B**3*CSV*CSV
      C**2=CON**1.5
      SG=C3?/A/B/SN1/SN4
      RT=C**2/A/B/SN1B/SN4B
      IF(ABS(AS-HPI).LE.1E-4)RT=RG*1E15
      IF(ABS(AS-HPI+PI).LE.1E-4)RT=RG*1E15
      CALL DIERT(DET,RG,RT)

      E=CEXP(CJ*TPI*(XO*SNT*CSP+YO*SNT*SNP+ZO*CST))
      CALL FIAL(AL,VSR,VHS)
      RGTP=RG
      RGTAU=RGTP*SNA*SNA
      FCORR=(0.0,0.0)
      FZH=(0.,0.)
      FTAUS=(0.0,0.0)
      DO 44 J=1,2
      FZH=FZH+DFT(1,J)*ALH(1,J)*CEXP(-AL(1,J)-CJ*TPI*SS)
      FTAUS=FTAUS+DFT(2,J)*ALH(2,J)*CEXP(-AL(2,J)-CJ*TPI*SS)
      FCORR=FCORR+DBAR(J)*DFT(1,J)*ALH(1,J)*CEXP(-AL(1,J)-CJ*TPI*SS)
   44 CONTINUE
      FTAUS=FTAUS*RGTP/RGTAU
      CSBETA=COS(BETAR)
      SNBETA=SIN(BETAR)
      HOPE=-CJ/2*CEXP(CJ*PI/6)
      ETH=E*SNBETA*SNT*FTAUS
      FOH=E*(CSBETA*SNT*FZH-SCN*FZH*CST*SNBETA+SNBETA*(-SGN)*FCORR*HOPE*
     *FACWRG*CST)
      RETURN
      END

      FUNCTION FUNI(VR)
      COMMON/BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      FUNI=1/SQRT(A*A*SIN(VR)*SIN(VR)+B*B*COS(VR)*COS(VR))
      RETURN
      END

      SUBROUTINE FIAL(AL,VL,VU)
      COMPLEX AL(2,2),CE,CJ
      DIMENSION AQ(2,2),Q(2,2)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      COMMON /GTO/BS,MODE,IO,IOTX
      DATA Q/1.,1.1679,2.3381074l,3.2482,4.087949441
      DATA AQ/.53566,.761210A2,-.41902,-.88311137/
      DATA PI,TPI,DPR/3.14159265,6.2831853,57.29577951
      CJ=(0.,1.)
      CE=CEXP(CJ*PI/6)*PI**(1./3)
      T=ABS(SIN(AS-PI/2)/SIN(AS))/2
      IO=1
      CALL GINT(VL,VU,QG1)
      IO=2
      CALL GINT(VL,VU,QG2)
      DO 1 I=1,MODE
      AL(1,I)=Q(1,I)*CE*QG1+QG2*(Q(1,I)*Q(1,I)/60+(.1-T*T)/Q(1,I))/CE
      AL(2,I)=Q(2,I)*CE*QG1+QG2*Q(2,I)*Q(2,I)/60/CE
    1 CONTINUE
      RETURN
      END

      FUNCTION FUN(VP)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      X=A*SIN(VP)
      Y=B*COS(VP)
      FUN=SQRT(X*X+Y*Y)
      RETURN
      END

      SUBROUTINE LAUNCH(ALH,RG,RT)
      COMPLEX ALH(2,2),DFT(2,2),C12,B,CJ
      DIMENSION AQ(2,2)
      DATA AQ/.53566,.761210A2,-.41902,-.88311137/
      DATA PI/3.14159265/
      CALL DIERT(DET,RG,RT)
      CJ=(0.,1.)
      C12=CEXP(CJ*PI/12)
      A=(PI*RG)**(-1./3)
      DO 1 I=1,2
      B=PI*C12*A
      IF(I.EQ.2)B=PI*A*A/C12
      DO 1 M=1,2
      ALH(I,M)=B*DFT(I,M)*AQ(I,M)
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE CINT(VL,VU,Y)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      COMMON /GTD/AS,MODE,ID,IDIX
      EXTERNAL FCT
      Y=0
      IF(ABS(VU-VL).LE.1E-5) GO TO 2
      CALL DQG72(VL,VU,FCT,YY)
      Y=Y+YY
      IF(IDIX.EQ.2)Y=-Y
    2 RETURN
      END


      FUNCTION FCT(X)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      COMMON /GTD/AS,MODE,ID,IDIX
      DATA PI,TPI,DR/3.14159265,6.2831853,57.29577951/
      A2=A*A
      B2=B*B
      SNA=ABS(SIN(AS))
      SN=SIN(X)
      CS=COS(X)
      Q=(A2*B2*SNA)**(1./3)
      F=SQRT(A2*SN*SN+B2*CS*CS)
      IF(ID.EQ.3) GO TO 3
      IF(ID.EQ.2) GO TO 2
      FCT=Q/F
      RETURN
    2 FCT=Q*Q*SNA/F/F/F
    3 FCT=F
      RETURN
      END


      SUBROUTINE FINT(ANS,VIR,VFR)
      EXTERNAL FUN
      IF(ABS(VIR-VFR).LT.1E-5)GO TO 1
      CALL DQG32(VIR,VFR,FUN,ANS)
      IF(ANS.LT.0)ANS=-ANS
      RETURN
    1 ANS=0
      RETURN
      END


      SUBROUTINE FKARG(SKWIG,ALR,VIR,VFR)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      EXTERNAL FUNT
      IF(ABS(VIR-VFR).LT.1E-5) GO TO 1
      SKWIG=(3.14159265*ABS(SIN(ALR)))**(1./3)
      SKWIG=SKWIG*(A*B)**(2./3)
      CALL DQG32(VIR,VFR,FUNT,ANS)
      SKWIG=SKWIG*ANS
      IF(SKWIG.LT.0)SKWIG=-SKWIG
      RETURN
    1 SKWIG=0
      RETURN
      END

      SUBROUTINE UNITV(UN,UB,VS,ALR)
      DIMENSION UN(3),UB(3)
      COMMON /BLKABF/A,B,VSR,PHSR,ZS,JANT,SGN
      X=A*SIN(VS)
      Y=B*COS(VS)
      SRXY=SQRT(Y*X+Y*Y)
      UN(1)=Y/SRXY
      UN(2)=X/SRXY
      UN(3)=0
      UB(1)=X*COS(ALR)/SRXY
      UB(2)=-Y*COS(ALR)/SRXY
      UB(3)=SIN(ALR)
      RETURN
      END


      SUBROUTINE DIFRT(DFT,RG,RT)
      COMPLEX DFT(2,2),C12,CJ,A,B,C,D,E,CJ
      DIMENSION AO(2,2),C(2,2)
      DATA C/1.,1.01879,2.3381074,3.2462,4.087944/
      DATA AO/.57566,..701210 2,-.41902,-.80311137/
      DATA PI/3.14159265/
      CJ=(0.,1.0)
      C12=CEXP(-CJ*PI/12)
      C2=CEXP(-CJ*PI/2)
      RS=RT
```

```
      VF=VFD(N)
      PF=PFD(N)
      IF(X.LT.0) VF=VFL(N)
      IF(X.LT.0) PF=PFL(N)
      C=CL(N)*(X-XF(N)*SGN)
      C=1/C
      VG=VG+C*VF
      PG=PG+C*PF
    4 CONTINUE
      VG=VG*T
      PG=PG*T
      G=VG*CEXP(CJ*PG*PI/180)
      RETURN
 1000 VG=VFD(N)
      PG=PFD(N)
      IF(X.LT.0) VG=VFL(N)
      IF(X.LT.0) PG=PFL(N)
      G=VG*CEXP(CJ*PG*PI/180)
      RETURN
 2000 G=SQRT(X)
      RETURN
 3000 G=CJ*(2-CJ)/2/X**2
      RETURN
      END

      COMPLEX FUNCTION G(X)
      DIMENSION XF(10),CL(10),VFL(10),VFD(10),PFL(10),PFD(10)
      COMPLEX CJ
      DATA IN/0/
      DATA XF/0.,.25,.5,1.,1.5,2.,2.5,3.,3.5,4.5/
      DATA VFD/1.3994,1.232,1.0591,.7382,.4881,.3193,.2825,.13,.0835,
     2.0537/
      DATA PFD/0.,4.883,11.36,26.64,42.56,57.98,72.9,87.57,182.17,
     2116.7*/
      DATA VFL/1.3994,1.552,1.682,1.861,1.948,1.982,1.994,1.998,1.999,
     21.9995/
      DATA PFL/0.,-2.864,-3.9,-7.67,-2.42,-1.45,-.85,-.52,-.33,-.22/
      CJ=(0.,1.)
      PI=3.14159265
      SGN=SIGN(1.,X)
      IF(X.GT.4) GO TO 2000
      IF(X.LT.-2.5) GO TO 3000
      IF(IN.NE.0) GO TO 1
      DO 2 I=1,10
      CL(I)=1
      DO 2 N=1,10
      DX=XF(I)-XF(N)
      IF(ABS(DX).LT.1E-6) DX=1
      CL(I)=CL(I)*DX
    2 CONTINUE
      IN=1
    1 CONTINUE
      T=1
      DO 3 N=1,10
      DX=X-XF(N)*SGN
      IF(ABS(DX).LT.1E-6) GO TO 1000
      T=T*DX
    3 CONTINUE
      VG=0
      PG=0
      DO 4 N=1,10
      VF=VFD(N)
      PF=PFD(N)
      IF(X.LT.0) VF= VFL(N)
      IF(X.LT.0) PF=PFL(N)
      C=PL(N)*(X-XF(N)*SGN)
      C=SGN/C
      VG=VG+C*VF
      PG=PG+C*PF
    4 CONTINUE
      VG=VG*T
      PG=PG*T
      G=VG*CEXP(CJ*PG*PI/180)
      RETURN
 1000 VG=VFD(N)
      PG=PFD(N)
      IF(X.LT.0) VG=VFL(N)
      IF(X.LT.0) PG=PFL(N)
      G=VG*CEXP(CJ*PG*PI/180)
      RETURN
 2000 G=1.5775*CEXP(CJ*.5*X)*EXP(-.8823*X)
      RETURN

 3000 G=2*CJ/X**3-2/X**6
      GG=CJ*2*7.228/X**9
      RETURN
      END
```

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

A MULTIMODE FIBER-OPTIC COUPLER WITH LOW INSERTION LOSS

Prepared by:	Marvin D. Drake, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Electrical Engineering, Florida Institute of Technology
Assignment:	
(Air Force Base)	Hanscom AFB
(Division)	Deputy for Development Plans
USAF Research Colleague:	Donald B. Brick, Ph.D.
Date:	August 26, 1977
Contract No:	AF4460-EES

A MULTIMODE FIBER-OPTIC COUPLER WITH LOW INSERTION LOSS

by

Marvin D. Drake

ABSTRACT

A fiber-optic coupler is described which is simple to fabricate, is directional, leaves the fiber intact, and has an insertion loss of < .05 dB. The coupler utilizes the energy contained in the leaky modes of the fiber. The coupling ratio may be varied by changing the type of light source and input conditions. Possible applications are to optical data buss links of up to a few hundred meters, or as power or mode monitors.

ACKNOWLEDGMENTS

I am sincerely grateful to the Air Force Office of Scientific Research for sponsoring this research program, and to the American Society of Engineering Education and Mr. Fred O'Brien of Auburn University for an excellent and active administration of the program.

I particularly wish to thank Dr. Donald B. Brick of the Deputy for Development Plans, Electronic Systems Division, and Mr. Charles Kleekamp and Dr. Bruce Metcalf of the MITRE Corporation for their help and encouragement throughout the research. Dr. Brick provided an excellent opportunity to pursue research topics and the means to obtain information on the state-of-the-art. Messrs Kleekamp and Metcalf shared the results of research on fiber-optics and, in particular, their research on connectors. They also shared their laboratory facilities with me which made the experimental work possible.

Many other people from ESD and MITRE have aided, advised and assisted me through this program, and I thank them for helping make this a most enjoyable and productive summer.

A Multimode Fiber-Optic Coupler with Low Insertion Loss

I. Introduction

The coupling of energy from multimode optical fibers usually involves breaking the fiber and inserting a coupler,[1,2] fusing the fiber with another fiber[3] or stripping the cladding from the fiber and cementing the fibers together.[4] This manipulation of the fiber introduces losses, from a few tenths to several dB per coupler, which limits the number of couplers that may be placed on the fiber without resorting to a repeater.[1] An alternate means of coupling energy from the fiber is to utilize the leaky modes present on the fiber. These modes may propagate and contain significant energy for several hundred meters[5]. In addition, localized leaky modes may be created by bending or stressing the fiber, and the energy from these modes coupled out.[6,7]

The purpose of this report is to describe a multimode fiber-optic directional coupler which has essentially no insertion loss since it utilizes energy usually lost from the fiber. The coupler does not require the fiber to be broken or fused, or the cladding to be stripped from the core. The coupler may be used with step or graded index fiber. The coupling ratio is calculated in Section III using a simplified theory of multimode optical coupling proposed by Ogawa and modified for the present case.[4] The experimental results presented in Section IV show a forward coupling ratio much greater than the theoretical value. A possible explanation of the larger experimental coupling ratio, based on perturbation of the modes of the fiber, is discussed in Section V. Possible applications of the coupler are discussed in Section VI, and recommendations for further research in Section VII.

II. Description of the Coupler

The coupler, as shown in Fig. 1, consists of three optical fibers with short sections of their protective jackets removed. The fibers are precisely held in contact with each other using three aligning rods. The three aligning rods are held together with heat shrink tubing. The coupler can be assembled with the two short fibers passing completely through the coupler, as shown in Fig. 1, or the bared ends of a short fibers inserted into the coupler from each end.

The position of the fibers in the alignment rods is shown in Figs. 2 and 3. The diameter of the alignment rods, is related to the diameter of the optical fibers by[8]

$$D_R = (5+2\sqrt{6}) \, d_f \qquad (1)$$

where: D_R = diameter of alignment rod

d_f = outer diameter of optical fiber

The diameter of the alignment rods used for the couplers was made slightly smaller than the value calculated above to ensure that the fibers would be retained in the coupler. The optical fibers used in the experiments were nominally 0.125 mm in diameter which required a 1.237 mm diameter alignment rod. The rods actually used were nominally 1.22 mm in diameter, and 25 mm long.

Both step index and graded index fibers were used in the experiments. The parameters of the fibers are listed in Table 1.

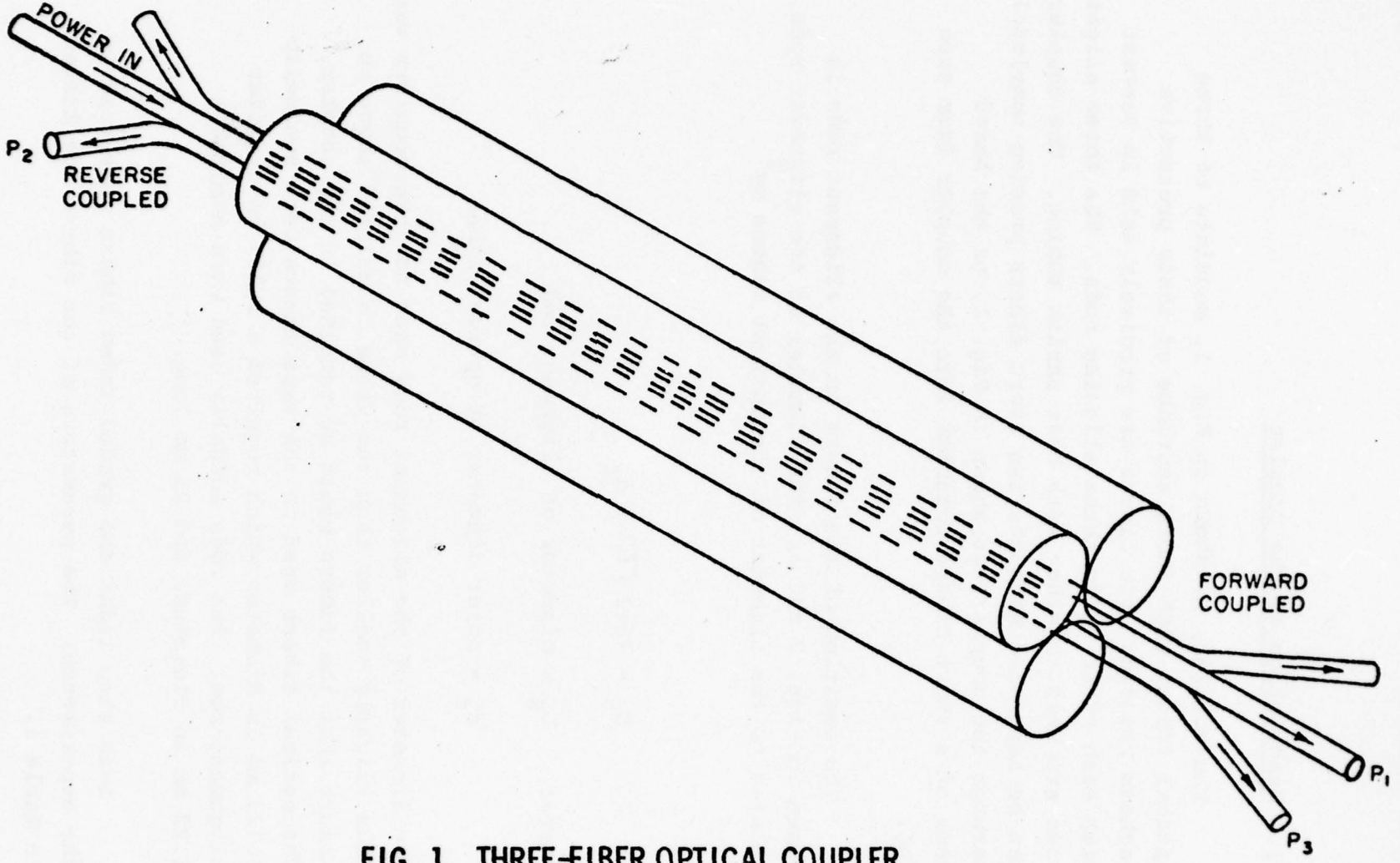


FIG. 1. THREE-FIBER OPTICAL COUPLER

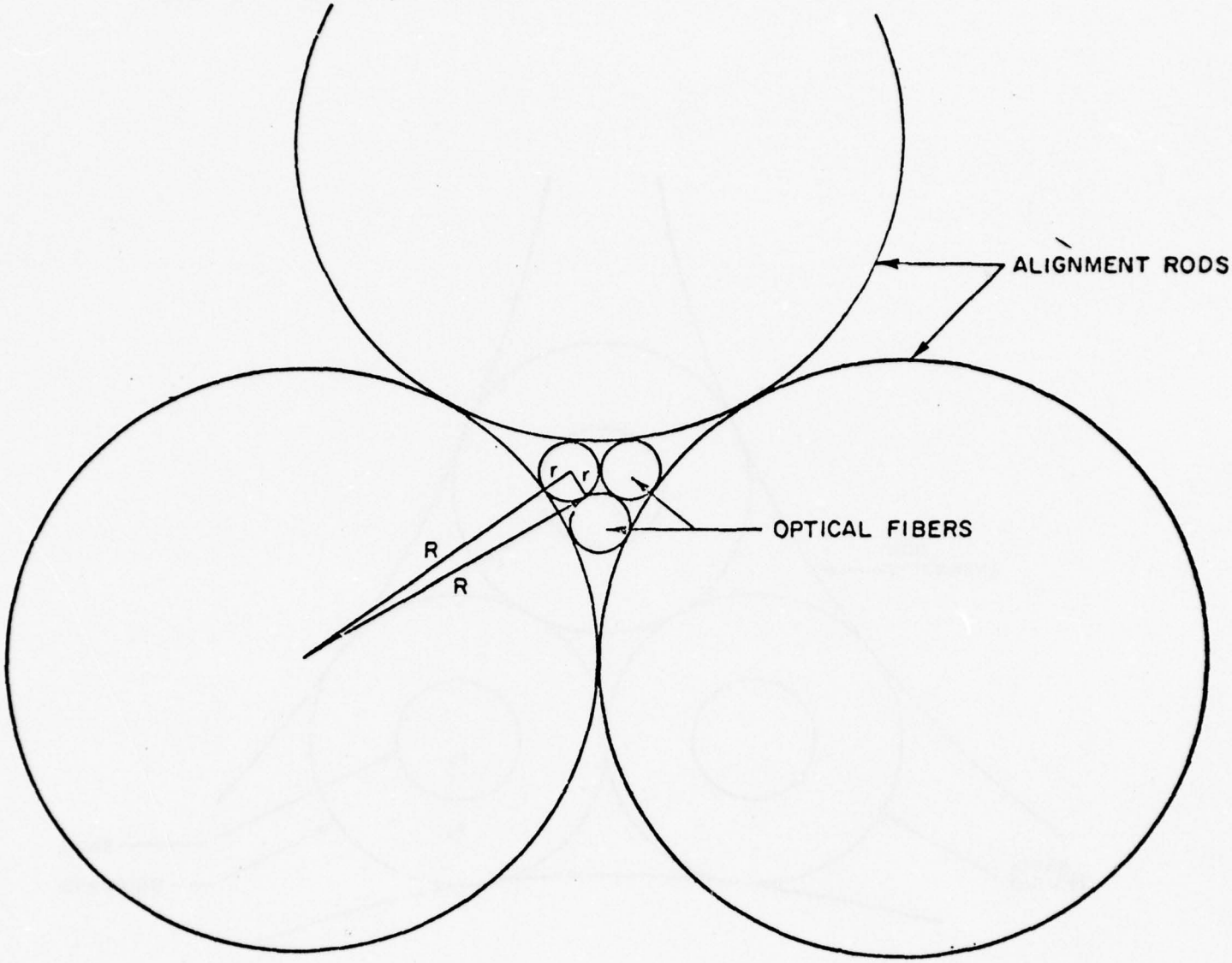


FIG. 2. CROSS-SECTION OF THREE FIBER COUPLER SHOWING GEOMETRY OF FIBERS AND ALIGNMENT RODS.

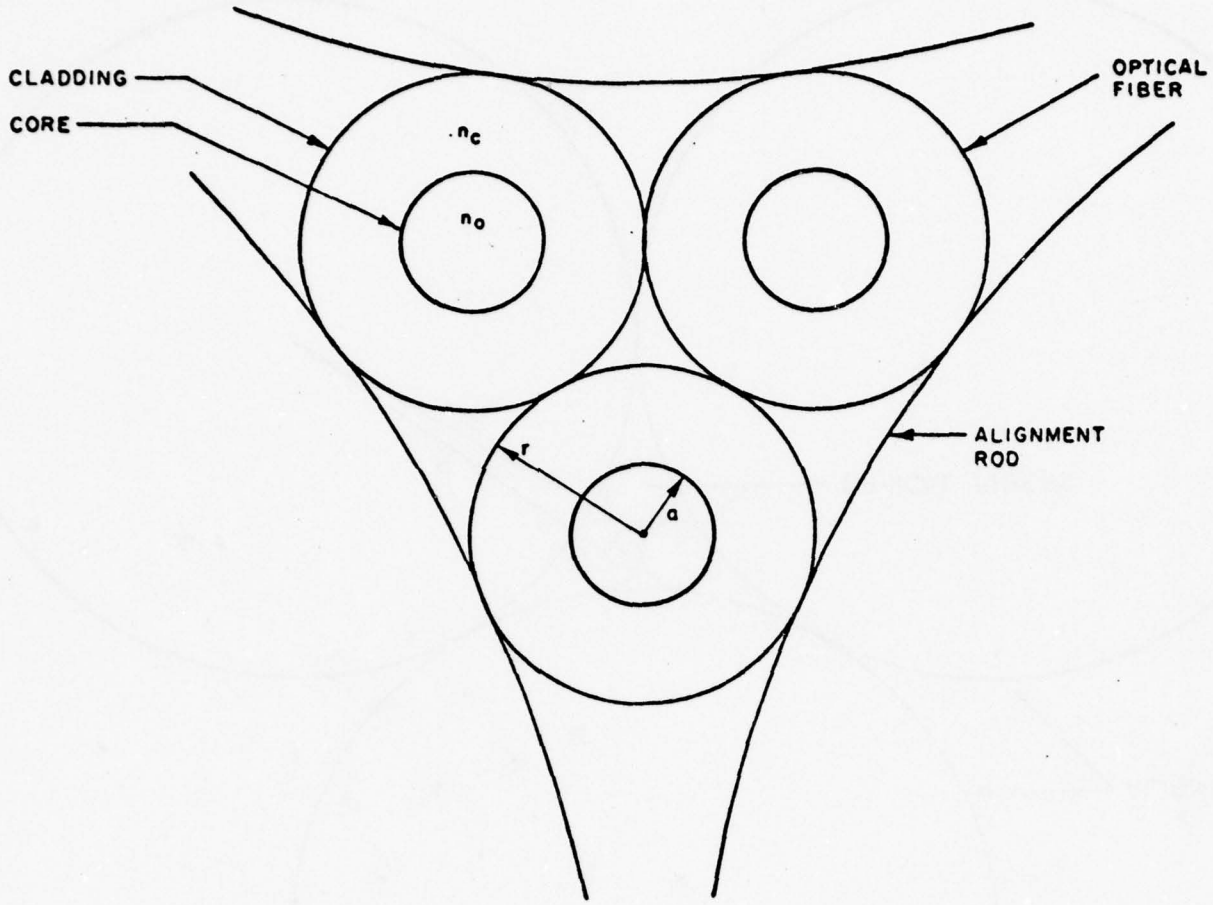


FIG. 3. DETAIL OF CROSS-SECTION OF THREE-FIBER COUPLER.

Table I

FIBER	STEP INDEX	GRADED INDEX
Manufacturer	ITT[9]	ITT[9]
Type No.	GS-02-12	GG-02-12
Fiber Core Diameter	50μm	50μm
Fiber Outer Diameter	125μm	125μm
Core Index of Refraction, n_o	1.48	1.48
Numerical Aperture	.25	.25
Loss (dB/km)	6.5	5.9

Several different lengths of input fiber (fiber #1) were used varying from 0.4m to 240m. The output fibers (#2 and #3) were usually short lengths (0.5 to 1m). The protective jacket was removed from fiber #1 about 0.5 m from the end of the fiber. The protective jacket was also removed from the ends of all of the fibers and the ends cleaved.[10]

Two sources were used in the experiments: a 1.5mw He-Ne laser and a 600 μW DH stripe geometry (50μm wide x 0.125 μm thick) LED.[11] The 1.5mm diameter He-Ne laser beam was coupled to the cleaved fiber end with either a 5X/0.10NA, 10x/0.30NA or 60x/0.80NA microscope objective. The LED was coupled to the input fiber by placing the cleaved face of the fiber as close as possible to the diode face. In some of the experiments, coupling of the LED was enhanced by forming a lens on the end of the fiber using an oxyacetylene torch.[12, 13]

III. Theoretical Derivation of Coupling Ratio

In this device, energy is coupled from the modes of the input fiber #1 into the modes of fibers #2 and #3. The types of fibers used are weakly guiding fibers, with $\Delta = (n_o - n_c)/n_o \ll 1$ where n_o and n_c are the index of refraction for the core and cladding respectively. In this case, Gloge[14] has given the transverse electromagnetic field components as:

$$E_y = H_x \begin{bmatrix} Z_o/n_o \\ Z_o/n_c \end{bmatrix} = E_\ell \begin{Bmatrix} J_\ell(ur/a)/J_\ell(u) \\ K_\ell(wr/a)/K_\ell(w) \end{Bmatrix} \begin{Bmatrix} \cos \ell\phi \\ \text{or} \\ \sin \ell\phi \end{Bmatrix} \begin{matrix} 0 \leq r \leq a \\ \\ r \geq a \end{matrix} \qquad (2)$$

where: $Z_o = \sqrt{\mu_o/\epsilon}$

E_ℓ is the electric field strength at the core-cladding interface and is given by

$$E_\ell = \left(\frac{\mu_o}{\epsilon}\right)^{1/4} \left(\sqrt{\frac{2}{\pi n_o}}\right) \frac{u K_\ell(w)}{a\left[v\sqrt{K_{\ell-1}(w) K_{\ell+1}(w)}\right]}, \quad \ell \neq 0 \qquad (3)$$

$$= \left(\frac{\mu_o}{\epsilon}\right)^{1/4} \left(\frac{1}{\sqrt{\pi n_o}}\right) \frac{u K_o(w)}{a v K_1(w)}, \quad \ell = 0$$

and $u = a(k^2 n_o^2 - \beta^2)^{\frac{1}{2}}$

$w = a(\beta^2 - k^2 n_c^2)^{\frac{1}{2}}$

$v^2 = u^2 + w^2 = a^2 k^2 (n_o^2 - n_c^2)$

β = propagation constant $\cong n_o k\left[1+\Delta - \Delta\,(u^2/w^2)\right]$

$k = 2\pi/\lambda$

r = radius from center of the fiber

$J_\ell(ur/a)$ = Bessel function of order ℓ

$K_\ell(wr/a)$ = modified Hankel function

For this formulation, $E_x = H_y = 0$ and the longitudinal field components, E_z and H_z are $(\Delta)^{\frac{1}{2}}$ in magnitude below E_y and H_x and are ignored. The region of interest is $r >> a$ i.e., the region outside the core and the cladding. Here the fields depend upon the form of $K_\ell(wr/a)$.

In our case, for most modes $w >> 1$, $w \cong v$ and $r >> a$, thus[14,15]

$$K_\ell(wr/a) \cong \sqrt{\frac{\pi a}{2wr}}\; e^{-\left(\frac{wr}{a}\right)} \tag{4}$$

and $$K_\ell(wr/a)/K_\ell(w) \cong \sqrt{\frac{a}{r}}\; e^{-w\left(\frac{r}{a}-1\right)}$$

thus $$E_y = H_x(Z_o/n_c) \cong E_\ell \sqrt{\frac{a}{r}}\; e^{-w\left(\frac{r}{a}-1\right)} \tag{5}$$

where $r > a$, $w >> 1$.

Thus since $w \gg 1$, the fields decay rapidly as the radius increases. However, the great number of modes propagating in this fiber, each contributing to the field far away from the fiber axis, will couple significant energy into an adjacent fiber. In addition, those few higher order modes which are close to cutoff ($w < < 1$) will carry most of their energy in the cladding of the fiber. In this case the fields can be approximated by [14, 15]

$$K_\ell (wr/a) \cong (\ell-1)!\, 2^{\ell-1} / (wr/a)^\ell , \qquad \begin{matrix} u = v \\ w << 1 \end{matrix} \qquad (6)$$

and $$E_y \cong E_\ell \, (a/r)^\ell$$

where $\ell \cong v$ for the highest order modes. A comparison of these for $v = 10$ is given in Table II and shows that the modes near cut-off do not decay as fast as those far from cut-off.

TABLE II

Comparison of $K_\ell(wr/a)/K_\ell(w)$ for higher order modes ($\ell \cong v$) close to and far from cut-off. (v=10)

r/a	$\sqrt{\frac{a}{r}}e^{-w(\frac{r}{a}-1)}$	$(a/r)^\ell$
1.0	1.0	1.0
1.2	0.1235	0.1615
1.4	0.0155	0.0346
1.6	1.960×10^{-3}	9.095×10^{-3}
1.8	0.250×10^{-3}	2.800×10^{-3}
2.0	$.0321\times10^{-3}$	0.9766×10^{-3}
2.2	4.142×10^{-6}	$0.376S\times10^{-3}$
2.4	0.5367×10^{-6}	0.1577×10^{-3}

A simplified theory for coupling between two identical multimode fibers has been formulated by Ogawa.[4] The model has the fibers completely surrounded by cladding material. This theory can be used as an upper bound for the coupler described in this report since the presence of the additional air-cladding interface should reduce the magnitude of the fields even more. Ogawa's theory is based on the coupling relations derived by Synder and Holntyre for coupling between two modes of different fibers.[16,17] The equations for coupling between a mode of fiber A and a mode of fiber B are:

$$dA(z)/d_z + j\beta_A A(z) = -jB(z)C_{AB}$$

$$dB(z)/d_z + j\beta_B B(z) = -jA(z)C_{BA} \qquad (7)$$

where B_A and B_B are the propagation constant for fibers A and B respectively and A(z) and B(z) are the amplitudes of the modes of fiber A and fiber B. C_{AB} and C_{AB} are the coupling coefficients from fiber A to B and from fiber B to A respectively.

$$C_{AB} = (\omega/2)\int_{\substack{\text{core area} \\ \text{of fiber B}}} E_o(n_o^2 - n_c^2) E_A E_B dS \qquad (8)$$

where E_A and E_B are the normalized electric fields in fibers A and B respectively. Using the same approximations for the modified Hankel functions as given earlier, Ogawa derives an average value of the coupling coefficient $|\overline{C_i}| = |C_{AiBi}| = |C_{BiAi}|$ between the $i^{\underline{th}}$ modes of the two fibers. Then $|\overline{C_i}|$ is the coupling/unit length of the fiber between the $i^{\underline{th}}$ mode of fiber A and the $i^{\underline{th}}$ mode of fiber B.

$$|\overline{C_i}| = \frac{2^{3/4}\,\Delta^{1/4}}{\sqrt{\pi k n_o a^3}} \left(\frac{i}{N}\right)\left(1 - \frac{i}{N}\right)^{1/4} \qquad (9)$$

$$\times \frac{\exp\left[-(2N-2i)^{1/2}\left(\frac{d}{a} - 2\right)\right]}{\sqrt{d/a}}$$

where: a = radius of core

$k = 2\pi/\lambda$

n_o = refractive index of core

$\Delta = 1 - \frac{n_c}{n_o}$

d=2R = distance between centers of the two fibers

N = total number of modes in the fiber

i = no. of the mode $\cong l^2/2 >> 1.$

The power coupled from the $i^{\underline{th}}$ mode of fiber A to the $i^{\underline{th}}$ mode of fiber B is then

$$P_{A \to B} = \sin^2\left[|\overline{C_i}|\ z\right] \tag{10}$$

where z = coupling length

If we treat the mode number i as a continuous variable, and if all modes carry equal power, then the coupling efficiency is

$$\eta_f = \frac{P_{A \to B}}{P_{in}} \leq \frac{1}{N}\int_0^N \left[\sin^2\left(|\overline{C_i}|z\right) di\right] \tag{11}$$

But for the present case where d/2a=2.5 and for coupling lengths of a few cm, then

$$\sin^2(|\overline{C_i}|z) \cong |\overline{C_i}|^2 z^2$$

$$\eta_f = \frac{z^2 A^2}{d/2a}\int_0^1 t^2(1-t)^{\frac{1}{2}}\exp\left[-b(1-t)^{\frac{1}{2}}\right] dt \qquad (12)$$

where $t = i/N$

$$A = \frac{(2\Delta)^{1/4}}{\sqrt{\pi k n_o a^3}}$$

$$b = 4\sqrt{2N}\left(\frac{d}{2a} - 1\right)$$

letting $x = (1-t)^{\frac{1}{2}}$

then: $$\eta_f \leq \frac{2z^2A^2}{d/2a}\int_0^1 (x^6 - 2x^4 + x^2)\, e^{-bx} dx$$

$$\eta_f \leq \frac{2z^2A^2}{d/2a} F(b) \qquad (13)$$

where $F(b) = \left[\frac{2}{b^3} - \frac{48}{b^5} + \frac{720}{b^7}\right]$

$$-e^{-b\left[\frac{8}{b^3} + \frac{72}{b^4} + \frac{312}{b^5} + \frac{720}{b^6} + \frac{720}{b^7}\right]}$$

Evaluating this for the case of the coupler made with step index fibers with z=25mm and

n_o = 1.48

NA = 0.25

d_{core}= 50 μm

$d_{cladding}$ = 125 μm

a = 25μm

d/2a = 2.5

λ = 633 nm

then k = $2\pi/\lambda$ = 9.926×10^{6}

Δ = 1.42668×10^{-2}

N = 1,924

A = 4.840×10^{2}

b = 1.861×10^{2}

then: η_f = 3.63×10^{-5} = -44.4dB

Figure 4 shows η_f as a function of coupling length, z, and as a function of d/2a allowing for variations in d in the fiber.

The experimental results presented in the next section show a forward coupling ratio much greater than that calculated above. The coupling also varies with excitation conditions and the distance of the tap from the source.

IV. Experimental Results

Two experimental setups were used as shown in Figures 5a and b. In figure 5a, a He-Ne laser was focused on the cleaved end of the fiber using a microscope objective. In 5b, the cleaved fiber face was positioned within 20 μm of the face of the LED. When a lens was formed on the end of the fiber, the end of the fiber was positioned 50 to 100μm from the face of the LED.

The insertion loss of all the couplers built is stated as ~0dB since the power through the fiber did not change appreciably as the coupler was assembled. The power through was measured on the original untouched fiber and monitored as the protective coating was stripped away and as the coupler was assembled. For a typical power through the fiber of 1.0mW the change in power through was <.01mW thus the insertion loss was < 0.05dB.

For each coupler, the power through fiber #1 was measured, the power coupled in the forward direction in fiber #3 and in the reverse direction in fiber #2 (see Figure '). The forward and reverse coupling ratios η_f and η_r and the directivity, η_d were calculated from

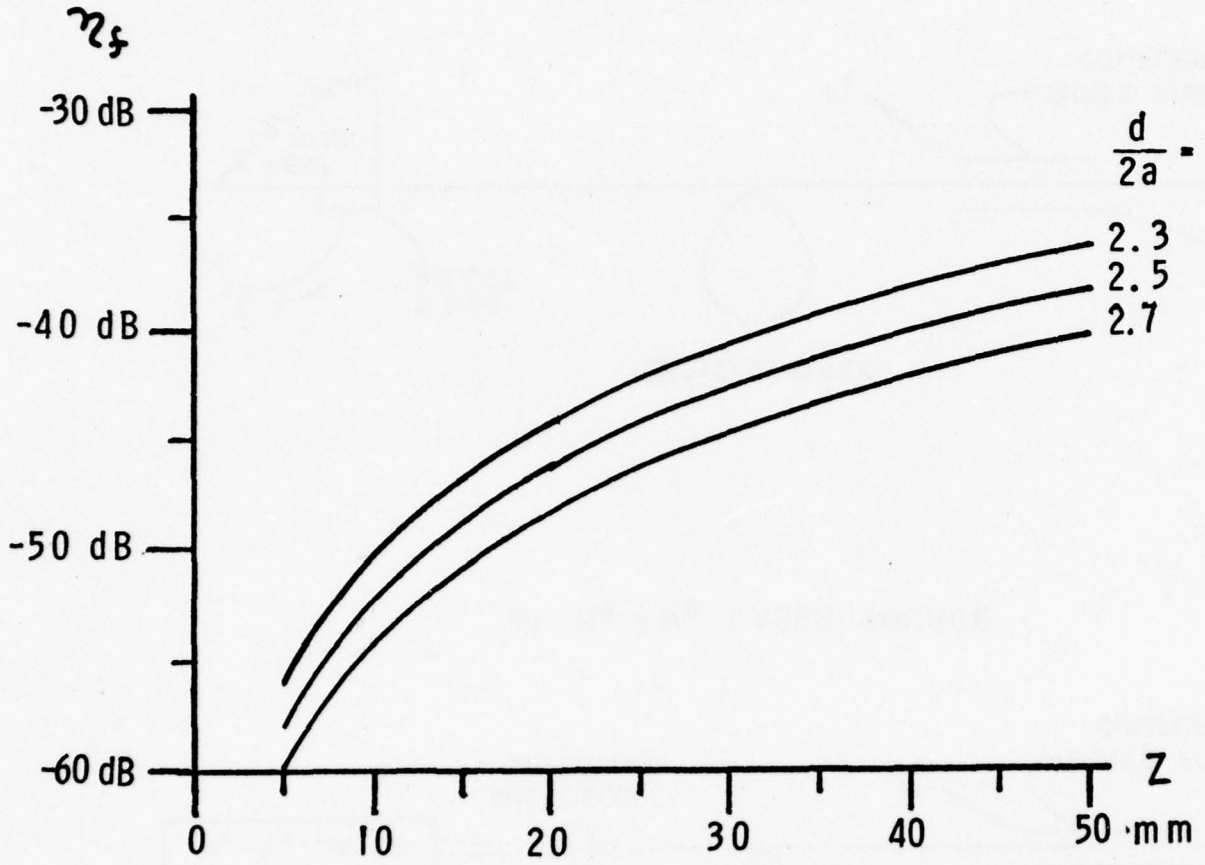
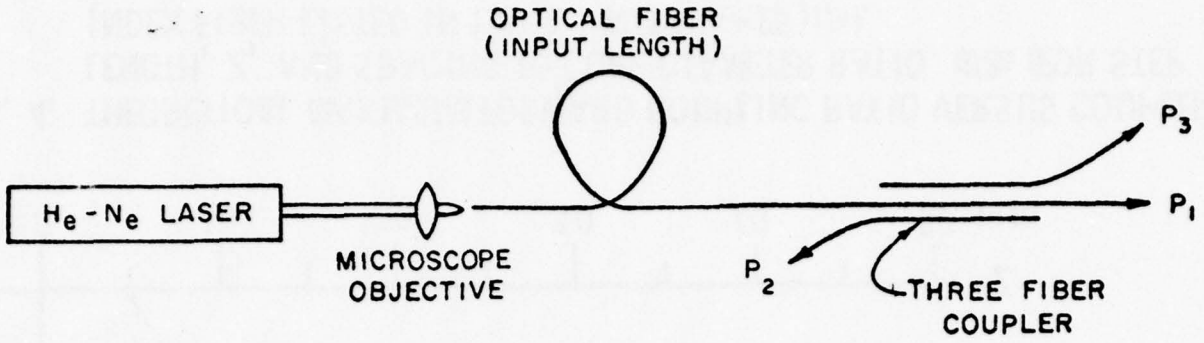
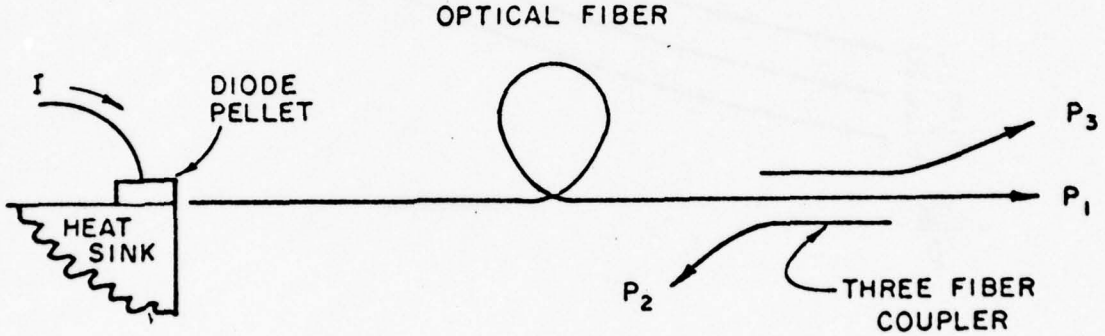


FIG. 4. THEORETICAL MAXIMUM FORWARD COUPLING RATIO VERSUS COUPLER LENGTH, Z, AND SPACING TO CORE DIAMETER RATIO, d/2a (FOR STEP INDEX FIBER LISTED IN TABLE I WITH $a = 25\ \mu m$).



a) H_e - N_e LASER SOURCE



b) GaAlAs EDGE EMITTER LED SOURCE

FIG. 5. EXPERIMENTAL SET-UP. a) WITH He-Ne laser source, b) WITH GaAlAs LED SOURCE (λ = 850 nm).

$$\eta_f = 10 \log (P_3/P_1)$$

$$\eta_r = 10 \log (P_2/P_1) \qquad (14)$$

$$\eta_d = 10 \log (P_3/P_2)$$

The experimental results for the step index fiber are listed in Table III for He-Ne laser excitation and in Table IV for GaAlAs LED excitation. The results for the graded index fiber with He-Ne excitation are given in Table V and in Table VI for GaAlAs LED excitation.

TABLE III

FIBER OPTIC COUPLER - EXPERIMENTAL RESULTS
STEP INDEX FIBER WITH He-Ne LASER SOURCE

Input Length	Input Lens	η_f (dB)	η_r (dB)	η_d (dB)	P_1 (μW)	Excitation (adjusted for maximum of:
0.7	none	-10.7	-29.0	18.2	19	P_1
0.7	5X/0.10NA	-25.4	-30.5	5.1	900	P_1
0.7	5X/0.10NA	- 6.6	-29.0	22.4	160	P_3
0.7	10X/0.30NA	-28.9	-30.2	1.3	940	P_1
0.7	10X/0.30NA	- 5.5	-27.8	22.3	155	P_3
0.7	60X/0.80NA	-20.2	-31.2	11.0	470	P_1
0.7	60X/0.80NA	- 7.6	-29.3	21.7	63	P_3

TABLE III (continued)

240	none	-35.6	-40.4	4.8	11	P_1 & P_3	
240	5X/0.10NA	-37.5	-47.5	10.0	505	P_1 & P_3	
240	10X/0.30NA	-37.8	-54.8	17.0	600	P_1 & P_3	
240	60X/0.80NA	-33.2	-50.2	17.0	315	P_1 & P_3	
350 + 0.7	10X/0.30NA	-24.0	-40.0	16.0	100	P_1 & P_3	Input fiber butt connected
350 + 0.7	60X/0.80NA	-24.4	-37.4	13.0	110	P_1 & P_3	to 350 m of fiber

TABLE IV

FIBER OPTIC COUPLER - EXPERIMENTAL RESULTS

STEP INDEX FIBER WITH GaAlAs LED SOURCE

Input Length (m)	Input Lens	η_f (dB)	η_r (dB)	η_d (dB)	P_1 (μw)
1.0	cleaved end	-20.8	-	-	155
1.0	flame lens	-21.8	-38.5	16.7	280
240	cleaved lens	-25.6	-36.4	10.8	87
240	flame lens	-27.6	-	-	270

TABLE V

FIBER OPTIC COUPLER - EXPERIMENTAL RESULTS

GRADED INDEX FIBER WITH He-NE LASER SOURCE

Input Length (m)	Input Lens	η_f (dB)	η_r (dB)	η_d (dB)	P_1 (μW)	Excitation adjusted for maximum of:
0.5	none	-24.1	-34.0	10.9	28	P_1 and P_3
0.5	5X/0.10NA	-35.2	-	-	1050	P_1
0.5	5X/0.10NA	-17.9	-32.3	14.4	220	P_3
0.5	10X/0.30NA	-38.6	-36.0	-	1000	P_1
0.5	10X/0.30NA	-18.0	-30.4	12.4	185	P_3
0.5	60X/0.80NA	-25.9	-34.0	8.1	430	P_3

TABLE V (continued)

100	5X/0.10NA	-36.3	-53.8	17.5	720	P_1
100	5X/0.10NA	-30.8	-51.6	20.8	290	P_3
100	10X/0.30NA	-40.1	-58.6	18.5	720	P_1
100	10X/0.30NA	-27.5	-43.5	16.0	180	P_3
100	60X/0.80NA	-40.6	-46.2	5.6	750	P_1
100	60X/0.80NA	-28.8	-47.2	18.5	420	P_3

TABLE VI

FIBER OPTIC COUPLER - EXPERIMENTAL RESULTS

GRADED INDEX FIBER WITH GaAlAs LED SOURCE

Input Length (m)	Input Lens	η_f (dB)	η_r (dB)	η_d (dB)	P_1 (μW)	
100	Cleaved end	-26.7	-39.7	13.0	94	P_1 maximum
100	flame lens	-25.4	-42.0	16.6	160	P_1 maximum
100	flame lens	-18.6	-35.8	17.2	38	P_3 maximum:

The data is summarized in Figure 6 where the range of the data for all excitation conditions is indicated. The maximum coupling values represent excitation for maximum P_3 while the minimum values represent maximum P_1. It should be noted that the maximum value for the coupled power could be doubled in each case by inserting fiber #2 in the forward coupling direction and detecting the power coupled into both fibers #2 and #3.

Finally, the presence of a connector in the fiber immediately before the coupler enhances the coupling ratio as shown by the data point at 350 m + 0.7 m in Figure 6.

V. Discussion

The data show that the experimental forward coupling ratio is far above the theoretical maximum value even at a distance of 240 meters from the source. However, the coupling ratio decreases toward the theoretical maximum as the distance from the source increases. The influence of the excitation conditions also decreases with distance from the source. These effects can be explained by the concept of a "spatial transient" and a "spatial steady state" of the mode distribution in the fiber.[5] At the input, the power coupled into the fiber is distributed into refracting leaky modes, tunneling leaky modes and bound modes. The refracting leaky modes die out within a few millimeters but the tunneling leaky modes may persist for hundreds of meters. After this distance the power in the fiber is carried mainly by bound modes which are the spatial steady state of the fiber. The discrepancy between the theory of Ogawa and the experimental results reported here can be explained by considering the variation in the modal power distribution (the optical spatial state) in the fiber with distance.

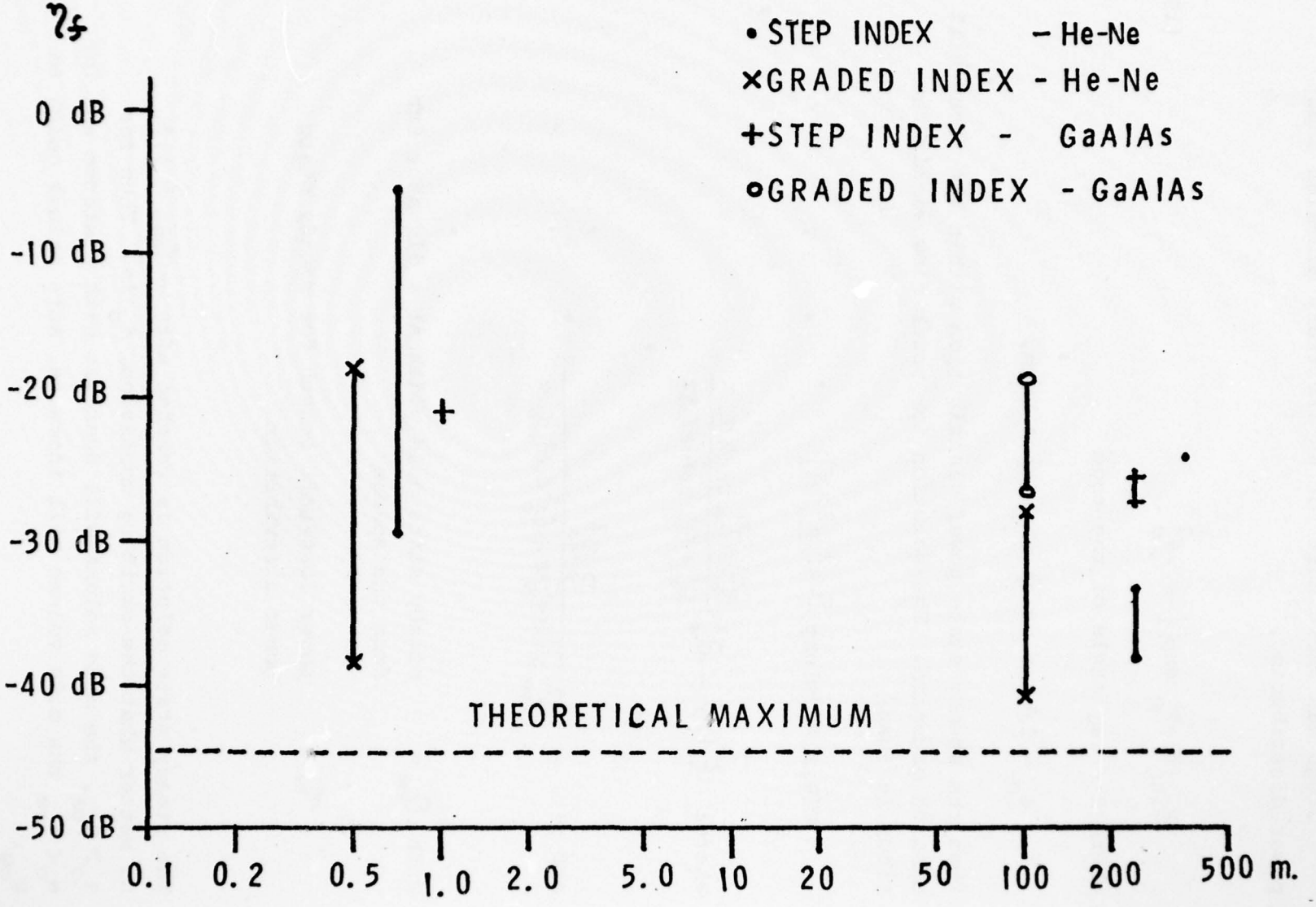


FIG. 6. SUMMARY OF FORWARD COUPLING RATIO VERSUS INPUT FIBER LENGTH DATA FOR ALL EXCITATION CONDITIONS.

Gloge has shown that for an arbitrary gaussian input power distribution,[18]

$$P_{in}(\theta)=P_o \exp\left[-\theta^2/\theta_o^2\right] \tag{15}$$

where θ = ray angle of the mode

θ_o = 1/e point of the input beam.

then the steady state power distribution within the fiber will also be gaussian. The solution for power flow within the fiber is then:

$$P(\theta,z)=f(z)\exp\left[-\theta^2/\underline{\theta}^2(z)\right] \tag{16}$$

where: $$\underline{\theta}^2(z) = \theta_\infty^2 \frac{\theta_o^2+\theta_\infty^2 \tanh \gamma_\infty z}{\theta_\infty^2+\theta_o^2 \tanh \gamma_\infty z}$$

and $$f(z) = \frac{P_o \theta_o^2}{\theta_\infty^2 \sinh \gamma_\infty z + \theta_o^2 \cosh \gamma_\infty z}$$

with θ_∞ = steady state mode width at a distance far from the source

γ_∞ = power loss/unit length for steady state power distribution

The steady state solution is reached after $\sim \gamma_\infty z \geq 1.5$, no matter what the exciting condition, θ_o is. Thus for $\theta_o > \theta_\infty$, the mode volume will decrease with distance and for $\theta_o < \theta_\infty$ the mode volume will increase, both values going to θ_∞.

The theory of Ogawa considered the case of equal power in each mode. However, this is not the case in the present coupler. Instead, if we use the distribution of Gloge above, and use

$P_i(z)$ = power in the $i^{\underline{th}}$ mode at a distance z from the source.

where i = mode number then:

$$P_i(z) = f(z)\ \exp\left[-\left(\frac{m\lambda}{4an_o}\right)^2 / \underline{\theta}^2(z)\right] \tag{17}$$

where m = mode group number:

The mode number, i for large m is $i \approx m^2$

Thus: $\left(\frac{m\lambda}{4an_o}\right)^2 \cong \frac{\pi^2\Delta}{4}\left(\frac{i}{N}\right) = \frac{\pi^2\Delta}{4}t$

and $$\eta_f \leqq \frac{z^2A^2}{d/2a}\int_0^1 t^2(1-t)^{1/2}\exp\left[-b(1-t)^{1/2}\right]f'(z)\exp\left[-\frac{\pi^2\Delta}{4}t/\underline{\theta}^2(z)\right]dt \tag{18}$$

This represents a much lower coupling coefficient than given by the theory of Ogawa <u>except</u> when $\theta_o >> \theta_{oo}$, and $\gamma_{oo} z << 1$.

In the latter case, the distribution tends toward that assumed by Ogawa. Greater coupling can be realized if the power in the higher order leaky modes can be increased. This can be done by varying the exciting conditions of the fiber or by perturbing the fiber. The former case only works for distances between the source and coupler of $z << \frac{1}{\gamma_{oo}}$. The perturbing of the fiber can include stressing the fiber to change the index of refraction distribution of the fiber,[6]. or changing the cladding configuration of the fiber during

manufacture. Stressing the fiber yields local mode conversion from bound and tunneling leaky modes to refracting leaky modes (or locally leaky modes).[7]

Changing the cladding configuration can change the amount of tunneling leaky modes present on the fiber and also the distance over which they persist. In particular the type of fiber known as a W-fiber or doubly clad fiber has the construction and index of refraction distribution shown in Figure 7. This type of fiber is characterized by a thin intermediate layer of lower index of refraction between the core and the cladding. The purpose of this layer is to confine the modes of the fiber more strongly within the core (i.e. in bound modes). However, Maeda and Yamada have recently shown that if this layer thickness is not carefully controlled ($\pm$ 0.5 μm), its may instead to leaky modes being excited which may persist to distances greater than 1 km.[19]

The type of optical source is also a factor in the amount of power present in the higher modes of the fiber. The data for the GaAlAs LED indicates a higher coupling ratio at longer lengths than for the He-Ne laser. This is to be expected since the LED may be considered as a lambertian source which excites more of the tunneling leaky modes than would a highly collinated source, such as the He-Ne laser. Further, injection lasers (ILD's) and LED's are rapidly becoming the source of choice for most fiber optic systems because of the small size and ease of coupling to the fiber. Thus, it may be expected that leaky modes will continue to be present on most fiber optic links in the future.

VI. Conclusions.

This report has described a fiber optic coupler which is simple to fabricate, is directional, leaves the fiber intact and thus has essentially zero insertion loss. The

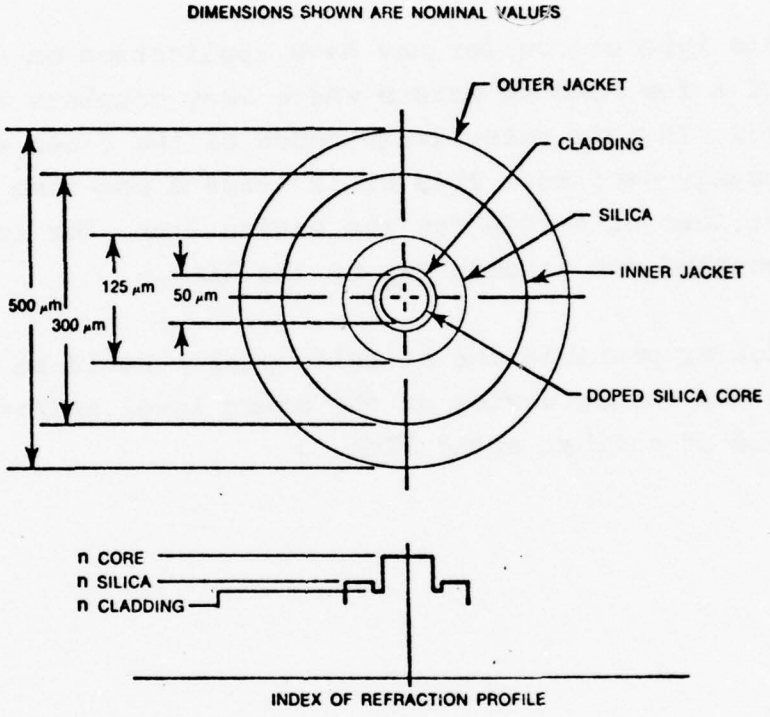


FIG. 7. DIMENSIONS AND INDEX OF REFRACTION PROFILE OF A W-TYPE FIBER (COURTESY OF ITT ELECTRO-OPTICAL PRODUCTS DIVISION).

coupler depends upon the energy contained in leaky modes of the fiber for its operation. The coupling ratio can be varied by varying the exciting conditions for distances up to 250 meters, and at distances greater than this by perturbing the fiber. The value of this coupler lies in that it utilizes power in the leaky modes which otherwise would be lost to radiation. Because of its dependence on input conditions, this type of coupler is well suited to use with ILD and LED sources.

This type of coupler may have application on data buss links of a few hundred meters where many couplers are required. In this case, leaky modes of the fiber would be deliberately excited. This would trade a one-time power thru-put loss of 6-10dB for the entire link, for an increased power coupled out at each tap on the link.

Another possible use of this coupler would be as a passive monitoring device of the power level and/or mode structure of a fiber optic link.

VII. Recommendations for Further Study

1. Measure coupling ratio on fibers from several manufacturers to determine applicability of this coupler.

2. Verify the theoretical variation of coupling ratio with length of the coupler.

3. Externally perturb the fiber statically or dynamically to vary the coupling ratio. Utilize a dynamic perturbation as a data input to the fiber.

4. Use the coupler to study the presence and decay of the spatial transient on fibers by placing the coupler at various positions on the fiber without breaking the fiber or changing the optical power through the fiber.

References

1. D.E. Altman and H.F. Taylor, "An Eight-Terminal Fiber Optics Data Bus Using Tee Couplers", Fiber and Integrated Optics, Vol. 1, No. 2 (1977) pp. 135-152.

2. F. Auracher and H.-H. Witte, "New Planar Optical Coupler for a Data Bus System with Single Multimode Fibers", Applied Optics, Vol. 16, No. 8 (August 1977), pp. 2195-2197.

3. B.S. Kawasaki and K.O. Hill, "Low-Loss Access Coupler for Multimode Optical Fiber Distribution Networks", Applied Optics, Vol. 16, No. 7 (July 1977) p. 1794-5.

4. K. Ogawa, "Simplified Theory of the Multimode Fiber Coupler", Bell Sys. Tech. J., Vol. 56, No. 5, (May-June 1977) pp. 729-745.

5. A.W. Snyder and Colin Pask, "Optical Fibre: Spatial Transient and Steady State", Optics Comm., Vol. 15, No. 2 (Oct 1975) pp 314-316.

6. P.D. Gianino and B. Bendow, "Refractive Index Changes in Optical Fibers Subject to Diametral Stress", RADC-TR-77-140, Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, New York 13441 (April 1977).

7. L. Jeunhomme and J.P. Pocholle, "Directional Coupler for Multimode Optical Fibers", Appl. Phys. Lett., Vol 29, No. 8 (15 Oct. 1976), pp. 485-487.

8. C. Kleekamp and B. Metcalf, MITRE Corporation, Bedford, MA Private Communication.

9. ITT, Electro-optical Products Division, Roanoke, VA 24019.

10. D. Gloge, P.W. Smith, D.L. Bisbee and E.L. Chinnock, "Optical Fiber End Preparation for Low-Loss Splices", Bell Syst. Tech. J., Vol 52, No. 9, (Nov. 1973) pp. 1579-1588.

11. Model IRE-103, Laser Diode Laboratories, Metuchen, N.J. 08840

12. D. Kato, "Light Coupling from a Stripe-Geometry GaAs Diode Laser into an Optical Fiber with Spherical End", J. Appl. Phys., Vol 44, No. 6 (June 1973) pp. 2756-8.

13. J.P. Wittke, M. Ettenberg and H. Kressel, "High Radiance LED for Single-Fiber Optical Links", RCA Review, Vol. 37 (June 1976) pp. 159-183.

14. D. Gloge, "Weakly Guiding Fibers", Applied Optics, Vol. 10 (1971), pp. 2252-2258.

15. A.W. Snyder, "Asymptotic Expressions for Eigenfunctions and Eigenvalues of a Dielectric or Optical Waveguide", IEEE Trans. Micro. Th. and Tech. MTT-17 (Dec. 1969) pp. 1130-1138.

16. A.W. Snyder, "Coupled-Mode Theory of Optical Fibers", J. Opt. Soc. Am., Vol. 62 (Nov 1972), pp. 1267-1277.

17. P.D. Holntyre and A.W. Snyder, "Power Transfer Between Optical Fibers, " J. Opt. Soc. Am., Vol. 63 (Dec 1973), pp. 1518-1527.

18. D. Gloge, "Optical Power Flow in Multimode Fibers," Bell Syst. Tech. J., Vol 51 (1972), pp. 1767-1783.

19. M. Maeda and S. Yamada, "Leaky Modes on W-Fibers: Mode Structure and Attenuation", Applied Optics, Vol. 16, No. 8 (Aug. 1977), pp. 2198-2203.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

ANALYSIS OF A DISCREPANCY IN TRAPPED PROTON FLUXES IN THE SOUTH ATLANTIC ANOMALY

Prepared by:	Thomas E. Morgan
Academic Rank:	Assistant Professor
Department and University:	Department of Earth Sciences State University of New York College at Oswego
Assignment:	L. G. Hanscom AFB Air Force Geophysics Laboratory Space Physics Division Magnetospheric Dynamics Branch
USAF Research Colleague:	Charles Dubs
Date:	July 29, 1977
Contract No.	F44620-75-C-0031

ANALYSIS OF A DISCREPANCY IN TRAPPED PROTON FLUXES IN THE SOUTH ATLANTIC ANOMALY

by

T. E. Morgan

ABSTRACT

S72-1 proton flux observations in the South Atlantic anomaly have been re-examined from both theoretical and observational points of view. A review of the techniques and assumptions used in reducing the flux data indicates that such techniques probably are not responsible for the discrepancy as evident in the fluxes in the H_{min} representation.

This preliminary analysis includes a proton trajectory computed by integration of the equation of motion using a model of the geomagnetic field. While it may be feasible to calculate theoretical proton trajectories in the region of the anomaly as a way toward checking some of the basic assumptions in the system of magnetic coordinates, further numerical testing is required before any definite conclusions can be drawn concerning the stability of the computations.

ACKNOWLEDGEMENT

I wish to thank the Air Force Systems Command for supporting this summer research project, and to acknowledge the ASEE and Auburn University, and in particular Mr. Fred O'Brien and Col. David England, for their efficient administration of the project.

Thanks are also due to John Howard and Adolph Jursa of the Air Force Geophysics Laboratory for providing the facilities which allowed this research to be carried out.

I would also like to express my deep appreciation to Charles Dubs for his willingness to serve as research colleague in addition to his regular duties, and who generously provided time for numerous helpful discussions.

Finally, I wish to acknowledge the help of Robert Filz and Ernest Holeman for providing the S72-1 data and computer programs which served as the basis for this project.

LIST OF FIGURES

LIST OF TABLES

Correction to Fig. 4: the curves should be labeled so that the uppermost is 6 MeV and the lowest is 36 MeV.

INTRODUCTION

Air Force satellite S72-1, launched in the latter part of 1972, followed a nearly circular polar orbit at an altitude of approximately 740 km. The satellite detected protons, most of which were apparently trapped and mirroring, over the South Atlantic. When transformed to the standard system of magnetic coordinates (L, B_m) or, equivalently, (L, H_{min}), a discrepancy is found such that flux data taken at high H_{min} generally exceed that taken at lower H_{min}. The flux difference at L = 1.15 and H_{min} = 200 km is about an order of magnitude and tends to decrease when going from L = 1.15 to L = 2.85, although it never disappears entirely. The data also suggest that the effect is greatest at the lowest proton energy measured (6 MeV), and tends to diminish when going to higher energy (36 MeV) for a given L shell. Figure 1 illustrates in a schematic manner the nature of the discrepancy.

Our initial objective was to examine the trapped motion of charged particles in the inner magnetosphere in terms of canonical perturbation theory, or a theory similar to this, and to apply this theory to the S72-1 observations. It has been customary to treat such data under the assumed validity of the adiabatic invariants, but in this case the resulting discrepancy appears to indicate that either an improved physical theory or a more careful treatment of the observational data may be necessary.

The observations of proton fluxes by S72-1 in the region of the South Atlantic anomaly are unique in that pitch angles have been measured to an accuracy on the order of 12^o. The transformation of counts recorded by the spacecraft detector into proton fluxes has been done in a fairly complex fashion, primarily in order to facilitate the physical interpretation of the data. We have examined the averaging processes used in reducing the data and, while the original data tapes are no longer available, two summaries of particle fluxes are contained on magnetic tapes. One tape contains a summary of pitch angle distributions sorted according to 1^o increments in latitude and longitude; a second tape contains pitch angle distributions sorted according to magnetic coordinates L and H_{min}.

Along with an examination of the data reduction process, we have initiated a preliminary theoretical attack on the problem. Instead of using canonical perturbation theory or expanding the first two adiabatic invariants to second order, as we had at first anticipated, we have begun computing proton trajectories in the region of the anomaly using the 1965 IGRF model of the geomagnetic field with 1975 coefficients. The purpose of the computer program is to solve the equation of motion of

a charged particle without using the guiding center approximation, and also without explicit dependence on the mirror equation (first adiabatic invariant) which has been used in reducing the particle fluxes in terms of H_{min}. In the following sections we indicate in detail the nature of the discrepancy and the results of our analysis.

SATELLITE AND ORBIT CHARACTERISTICS

Satellite S72-1 had an orbital inclination of 98.4°; its perigee and apogee altitudes were 729 km and 750 km, respectively. Given the spin rate of 12 rpm, a complete 180° pitch angle scan took 2.5 sec. All channels were read out simultaneously once every .08 sec, corresponding to a change in direction of 5.6° at the given spin rate. A preliminary analysis of the proton data and also discussion of the satellite characteristics is given by Filz et al. (1974).

The observed energy ranges and their midpoints are given in Table 1. The rigidity is included because of the way in which the trajectory tracing computer program requires the particle energy to be specified; the values given here were computed according to Shea et al. (1965).

The satellite spin axis is oriented normal to its orbit plane. The detector itself spins in the meridional plane and, given the 5.6° movement between sampling intervals, a single observation corresponds to an 18° range in pitch angle. The detector has excellent spatial resolution due to its 6° half-angle viewing cone, and thus provides a detailed description of the flux as a function of pitch angle (Holeman, 1973). Since the acceptance cone is normal to the spin axis, the satellite did not directly observe fluxes from the east-west direction. A sketch of the detector opening is shown in Fig. 2, and Fig. 3 illustrates the detector orientation in space.

Although the actual detector opening is circular with a 6° half-angle, the middle 5.6° seen in Fig. 2 is due to the satellite spin motion during one sampling period. The resulting effective opening (something between an ellipse and a rectangle) is approximated by a rectangular shape broken down into nine 2° strips for analyzing the incoming particles. The 2° strips are indicated by horizontal dashed lines in the diagram.

As shown in Fig. 3, the satellite orbit plane lies in the plane of the paper. The satellite spin axis is perpendicular to the paper at the point indicated. After the simultaneous readout of all channels, the satellite

spins 5.6° before the subsequent readout takes place. Thus, the observations are smeared over 5.6°. (The linear distance travelled by the satellite between two successive observation points is about 0.6 km).

A general summary of the reduction procedure is given by Holeman (1973); Appendix C of this report contains a series of contour plots for B, L, H_{min}, and proton flux as a way of displaying the proton and magnetic coordinate characteristics. A series of computer programs rewrote the data from the incoming tapes (in excess of 800) onto a set of more compact intermediate tapes. The latter were then combined with the ephemeris and magnetic field coordinates, and the results were separated into two groups: data recorded in the vicinity of the South Atlantic anomaly, with which we are concerned here, and data recorded over the poles.

We should note that the satellite did not carry a magnetometer, so that the magnetic field associated with a given position is determined from a geomagnetic field model. Given the satellite's latitude, longitude, and orientation, the geomagnetic field model provided the magnetic field strength B and the pitch angle. In this way the satellite can be said to have observed pitch angle distributions at different values of L and B over the South Atlantic.

PROTON FLUX DISCREPANCY

We turn now to a detailed discussion of Fig. 1 introduced earlier. For a given value of L and five energies, the satellite observed pitch angle distributions at different values of B. The pitch angle distributions tend to be charply peaked such that most of the flux occurs at or close to 90° (see Fig. 4). The value of B is therefore the mirror field strength B_m for the 90° pitch angle flux. In Fig. 1, the seven B_m values observed have been transformed to H_{min}, and the standardized set of H_{min} values is used to label each curve in place of B_m. This procedure accounts for the right-hand point of each curve. For other pitch angles at which protons are observed, B_m is determined by means of the mirror equation, and the corresponding H_{min} is calculated. The observed flux is then plotted against this value of H_{min}.

This procedure is equivalent to starting at a low value of B at a high altitude (for a given L shell), and then following the flux as it mirrors further down at B_m and H_{min} (at higher values of B and thus lower H values). The discrepancy can therefore be described as follows. We should be able to predict the same value of the mirroring flux further

down along the field line at the proper pitch angle based on its observed (local) value. Figure 1 shows that, for the same field line, different observed fluxes do not coincide when they are transformed to the same H_{min}. The vertical displacement of the curves from one another is a measure of the discrepancy.

DATA REDUCTION COMPUTATIONAL TECHNIQUES

As part of our attempt to understand the trend indicated in Fig. 1, we present a summary of the relevant computational techniques and assumptions contained in the series of FORTRAN programs written by E. Holeman for the purpose of treating S72-1 proton data.

Subroutine PACRC (XL, POUT, HM, ERG) calculates the 16 x 31 matrix POUT for each set of values for L, H_{min}, and E. We present the basic definitions of the information contained in the major matrix, followed by a more detailed discussion of several of the more important individual arrays.

POUT (1, J).....local pitch angle; the correct sequence is

J =	1,	2,	.....	,	30,	31.
for pitch angle	89°,	87°,	.....	,	31°,	15°
in the range	88-90°,	86-88°,	.....	,	30-32°,	0-30°;

POUT (2, J).....the H_{min} at which a proton of pitch angle J should mirror for a given set of input data;

POUT (3, J).....the observed local flux at pitch angle J;

POUT (4, J).....the omnidirectional differential flux;

POUT (5, J).....the 9-point averaged flux for J;

POUT (6, J).....the standard deviation associated with POUT (5, J);

POUT (7, J) } ....unfolded 9-point data;
POUT (15, J) }

POUT (16, J)....normalization factor.

1. POUT (2, J)

For a given L shell, a linear relation has been fitted to the field B as a function of minimum altitude H_{min} in the form

$$B = BEX \cdot H_{min} + BAX$$

or

$$B = BEX\ (H_{min} + BFX)$$

where

$$BFX = BAX/BEX .$$

Solving for H_{min}, we have:

$$H_{min} = \frac{B}{BEX} - BFX.$$

Here, BEX and BAX are the constants of the linear fit. Now, a particle observed at a given value of B and pitch angle α should mirror at

$$B_m = \frac{B}{\sin^2 \alpha} \quad \text{(mirror equation)}$$

which corresponds to some value of the minimum mirror point altitude $H_{mm} \leq H_{min}$. Here, H_{min} is defined in terms of McIlwain's B, L coordinate system as the minimum altitude encountered by following a line of constant B completely around a surface of constant L (Holeman, 1973, p. 4). In the same way, H_{mm} is the minimum mirror point altitude encountered by following a line of constant B_m completely around a surface of constant L. We can write:

$$H_{mm} = \frac{B_m}{BEX} - BFX.$$

Substituting for $B_m = \frac{B}{\sin^2 \alpha} = \frac{BEX\ (H_{min} + BFX)}{\sin^2 \alpha}$

we get

$$H_{mm} = \frac{H_{min} + BFX}{\sin^2 \alpha} - BFX.$$

H_{mm} is the value of POUT (2, J) in PACRC; the constants BEX and BAX are tabulated for each altitude range and are interpolated at a given L from subroutine INTER.

2. POUT (3, J)

The observed local flux at pitch angle J is given by

$$\text{POUT}(3, J) = \frac{\text{FXK1} \cdot \text{ICAD}}{\text{DE} \cdot \text{NCAD}} - \text{XBKG}$$

where i) FXK1 is a constant involving the geometrical factors of the detector assembly;

ii) ICAD is the number of counts for a given L, H_{min}, E;

iii) NCAD is the number of observations for a given L, H_{min}, E summed over a time period of five months;

iv) DE is the width of the energy bin, ΔE, and is used to derive the flux per unit MeV;

v) $\text{XBKG} = \frac{\text{FXK1} \cdot \text{ICAD}}{\text{DE} \cdot \text{NCAD}}$

is the background which is subtracted out at a given L, H_{min}, E. Here all counts having pitch angles in the range 0-30° are subtracted as background, leaving the 30-90° range for the observed local flux distribution.

3. POUT (4, J)

The omnidirectional differential flux (see Fig. 5) is stored in the block data subroutine PMODA. The matrices in PMODA are dimensioned 5x7x30. Interpolation is required only for a given H_{min}; no interpolation is required for the five E values or the 40 L values (not all L values are used since there are no data beyond L = 3.5). The interpolation is also exact for the seven standard H_{min} values: 250, 350, 450, 550, 625, 675, and 730 km. Although it may at first appear that POUT (4, J) has a pitch angle dependence, it is in fact the omnidirectional flux corresponding to (POUT (2, J), XL, ERG).

4. POUT (16, J)

At each pitch angle J in PACRC the following normalization factor is defined:

$$\text{POUT}(16, J) = \frac{\sum_{K=J-4}^{J+4} \text{POUT}(3, K)}{\sum_{K=J-4}^{J+4} \text{POUT}(4, K)} .$$

This is essentially the ratio of the observed unidirectional fluxes to the omnidirectional fluxes over 9 pitch angles having J as the central value of the pitch angle. Expanding POUT (16, J) in terms of the actual coding may make it easier to visualize the normalization process. The expanded coding is presented in Appendix A1, where we have set

$$\text{POUT } (3,K) = f_K \quad , \text{ unidirectional flux}$$

$$\text{POUT } (4,K) = F_K \quad , \text{ omnidirectional flux.}$$

In the expressions for POUT (16, J) we note that, because of the sharply defined pitch angle distributions (e.g., Fig. 4) in which most of the flux is found within a few degrees of mirroring, the numerator of each normalization factor will be most strongly influenced by the mirroring flux contribution.

5. POUT (5, J)

This array represents the final curves for the 9-point averaged flux for the J-th pitch angle:

$$\text{POUT } (5,J) = \frac{1}{9} \sum_{K=J-4}^{J+4} \text{POUT } (4,J) \cdot \text{POUT } (16,K).$$

Again, the expanded coding for this array is given in Appendix A2, with the following simplified notation:

$$\text{POUT } (4,J) = F_J \quad , \text{ omnidirectional flux}$$

$$\text{POUT } (16,K) = R_K \quad , \text{ normalization factor.}$$

We see that the averaged flux for the J-th pitch angle is equal to the omnidirectional flux for that H_{min} multiplied by the average of the J-th normalization factor and the eight neighboring factors.

6. POUT (6,1)

The first index of this array denotes these values as being the standard deviation; the second index denotes pitch angle (i.e., H_{min}). In the expanded coding given in Appendix A3 we show only the first and last members of this array, since the internal arrangement is similar to that of earlier expressions. The standard deviation is given in the form $\sigma = \left[\overline{x^2} - (\overline{x})^2\right]^{\frac{1}{2}}$.

For a final display of coding we illustrate the composition of the 9-point curves before averaging. These unfolded 9-point data are given in Appendix A4. Each 9-point curve, centered around the omnidirectional flux for that H_{min}, is multiplied by the same normalization factor.

PITCH ANGLE ANALYSIS

In view of the averaging processes which have been employed in ultimately presenting the data in Fig. 1, we have sought to go directly to the pitch angle distributions as a function of B and L. We computed H_{min} corresponding to B using the linear relationship presented earlier. This was done for all pitch angles at which particles were observed and, except for the mirroring flux, involved the use of the mirror equation. We then searched for the pitch angle in the observed distribution for which $H_{min} \cong 300$ km and, following Fischer et al. (1977), we denoted this pitch angle as defining the "loss cone," θ_{LC}. Generally, proton fluxes were observed at pitch angles greater than θ_{LC}. We denote the smallest pitch angle at which protons were observed by the symbol θ_s. The results are summarized in Tables IIa, b, c, d, and e.

The difference between the pitch angle defining the loss cone and the minimum pitch angle for which proton flux is observed is given by $\Delta\theta = \theta_{LC} - \theta_s$; ΔB is the amount by which the field corresponding to θ_s (i.e., B_s) exceeds the field corresponding to $H_{mm} \cong 300$ km. Here we define B_s as $B/\sin^2\theta_s$. We should note that the value of 300 km for H_{mm} is not a sharp cutoff. If we had chosen H_{mm} = 200 km instead, the values for $\Delta\theta$ and ΔB would be smaller. Besides pitch angle changes, another possibility is that the magnetic field may have been decreased in comparison with the value of the field predicted by the mirror equation.

THEORETICAL PROTON TRAJECTORY CALCULATIONS

A computer program, described in Shea et al. (1965) and Shea et al. (1976), with slight modifications was used to compute proton trajectories in the region of the South Atlantic anomaly. As input parameters the program requires rigidity, latitude, longitude, and altitude; initial pitch angle is determined by the choice of zenith and azimuth angles. The program utilizes a fourth-order Runge-Kutte technique to integrate the equation of motion of a proton in the geomagnetic field:

$$m\frac{d^2\vec{R}}{dt^2} = \frac{e}{c}\left(\frac{d\vec{R}}{dt} \times \vec{B}\right)$$

where $\vec{R}$ = position vector
$\vec{B}$ = vector magnetic induction in gauss
m = relativistic mass
c = speed of light
e = charge in electrostatic units

Although the program was originally designed to compute cosmic ray trajectories, we have attempted to use it in following a proton as it drfits in longitude. Figure 6 shows the northern and southern mirror point traces for a 6 MeV proton as it drifts across the anomaly region. The points making up the curve were selected as the lowest altitude reached by the particle in the gyration containing the mirror point. This should be a useful way to check the minimum mirror point altitude independently of any assumptions regarding the magnetic coordinates and the linear relation between B and H_{min}. However, the amount of computer time required for a proton to drift across the anomaly is quite long and, depending on the time step chosen, may exceed several hours.

The proton traces shown in Fig. 6 took about two hours of computer time on a CDC 6600. The proton was started at an altitude of 740 km, and was given a latitude of -27.5°, and a longitude of -58.5°; the initial pitch angle was 89.56°. With the time step positive, the proton drifted west; we then reversed the time step by making it negative, thus obtaining the proton's trajectory east of the starting point. We are able to modify the time step such that a single gyration may take on the order of 50 or 100 integration steps or, such as Fig. 6, 200 integration steps. It is possible that by making the time step too small, significant roundoff errors may occur. Further numerical testing is required before any definite conclusions can be drawn concerning the stability of the computation.

SUMMARY AND CONCLUSIONS

The high resolution of proton fluxes in terms of pitch angle obtained by S72-1 in late 1972 and early 1973 have not been superseded by other observations, and therefore are worthy of further careful study. In analyzing the mirroring flux curves in Fig. 1, we have been led to a review of the data reduction techniques. Any reanalysis by a different (and as yet unspecified) reduction procedure would require the use of the L, H_{min} or latitude, longitude, pitch angle data tapes.

Theoretical proton trajectories in the region of the anomaly can be used to check the validity of Liouville's theorem as follows. A proton can be released in the computer simulation at a pitch angle which is different from 90° (our proton in Fig. 6 does not really qualify for this test). The proton would be released at 740 km, the spacecraft altitude, and as it mirrored it would seek a lower altitude in the anomaly. Since this is the region where the magnetic shells lie closest to the earth, the particle will drift such that its mirror point eventually again reaches 740 km

altitude. Then the flux at the injected pitch angle should be the same as the flux at the 90^{o} pitch angle at 740 km on both sides of the anomaly. This comparison should be made for a range of positions and pitch angles.

However, a more efficient computational technique would be one such as that described by Heckman and Brady (1966), in which a predictor-corrector method is used to continue the integration after a Runge-Kutta technique has computed the necessary starting values.

A decrease Δ B in the magnetic field may possibly help explain the proton flux discrepancy. Such a decrease was derived by Sugiura (1972), in which large field depressions on the order of 40 gammas were deduced in the equatorial region due to the presence of an equatorial current sheet in the magnetosphere. However, the region studied was beyond about 2.5 R_E, and so the results do not appear directly applicable to the L region covered by the S72-1 data. It is possible that, due to the high resolution in the pitch angle distribution, the model field used is not accurate enough for interpreting the observed proton fluxes in the region of the anomaly.

APPENDIX A1

$$POUT(16,1) = \frac{f_4 + f_3 + f_2 + f_1 + f_1 + f_2 + f_3 + f_4 + f_5}{F_4 + F_3 + F_2 + F_1 + F_1 + F_2 + F_3 + F_4 + F_5} = R_1$$

$$POUT(16,2) = \frac{f_3 + f_2 + f_1 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6}{F_3 + F_2 + F_1 + F_1 + F_2 + F_3 + F_4 + F_5 + F_6} = R_2$$

$$POUT(16,3) = \frac{f_2 + f_1 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7}{F_2 + F_1 + F_1 + F_2 + F_3 + F_4 + F_5 + F_6 + F_7} = R_3$$

$$POUT(16,4) = \frac{f_1 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 + f_8}{F_1 + F_1 + F_2 + F_3 + F_4 + F_5 + F_6 + F_7 + F_8} = R_4$$

$$\vdots$$

$$POUT(16,21) = \frac{f_{17} + f_{18} + f_{19} + f_{20} + f_{21} + f_{22} + f_{23} + f_{24} + f_{25}}{F_{17} + F_{18} + F_{19} + F_{20} + F_{21} + F_{22} + F_{23} + F_{24} + F_{25}} = R_{21}$$

$$\vdots$$

$$POUT(16,28) = \frac{f_{24} + f_{25} + f_{26} + f_{27} + f_{28} + f_{29} + f_{30} + f_{30} + f_{30}}{F_{24} + F_{25} + F_{26} + F_{27} + F_{28} + F_{29} + F_{30} + F_{30} + F_{30}} = R_{28}$$

$$POUT(16,29) = \frac{f_{25} + f_{26} + f_{27} + f_{28} + f_{29} + f_{30} + f_{30} + f_{30} + f_{30}}{F_{25} + F_{26} + F_{27} + F_{28} + F_{29} + F_{30} + F_{30} + F_{30} + F_{30}} = R_{29}$$

$$POUT(16,30) = \frac{f_{26} + f_{27} + f_{28} + f_{29} + f_{30} + f_{30} + f_{30} + f_{30} + f_{30}}{F_{26} + F_{27} + F_{28} + F_{29} + F_{30} + F_{30} + F_{30} + F_{30} + F_{30}} = R_{30}$$

APPENDIX A2

$$\text{POUT}(5,1) = \frac{1}{9} F_1 \left[R_4 + R_3 + R_2 + R_1 + \underset{\downarrow}{R_1} + R_2 + R_3 + R_4 + R_5 \right] = \bar{F}_1$$

$$\text{POUT}(5,2) = \frac{1}{9} F_2 \left[R_3 + R_2 + R_1 + R_1 + R_2 + R_3 + R_4 + R_5 + R_6 \right] = \bar{F}_2$$

$$\vdots$$

$$\text{POUT}(5,20) = \frac{1}{9} F_{20} \left[R_{16} + R_{17} + R_{18} + R_{19} + R_{20} + R_{21} + R_{22} + R_{23} + R_{24} \right] = \bar{F}_{20}$$

$$\vdots$$

$$\text{POUT}(5,26) = \frac{1}{9} F_{26} \left[R_{22} + R_{23} + R_{24} + R_{25} + R_{26} + R_{27} + R_{28} + R_{29} + R_{30} \right] = \bar{F}_{26}$$

$$\vdots$$

$$\text{POUT}(5,30) = \frac{1}{9} F_{30} \left[R_{26} + R_{27} + R_{28} + R_{29} + R_{30} + R_{30} + R_{30} + R_{30} + R_{30} \right] = \bar{F}_{30}$$

APPENDIX A3

$$\text{POUT}(6,1) = \left\{ \frac{1}{9} F_1^2 \left[R_4^2 + R_3^2 + R_2^2 + R_1^2 + \underset{\downarrow}{R_1^2} + R_2^2 + R_3^2 + R_4^2 + R_5^2 \right] - \frac{1}{81} F_1^2 \left[R_4 + R_3 + R_2 + R_1 + \underset{\downarrow}{R_1} + R_2 + R_3 + R_4 + R_5 \right]^2 \right\}^{\frac{1}{2}}$$

$$\vdots$$

$$\text{POUT}(6,30) = \left\{ \frac{1}{9} F_{30}^2 \left[R_{26}^2 + R_{27}^2 + R_{28}^2 + R_{29}^2 + R_{30}^2 + R_{30}^2 + R_{30}^2 + R_{30}^2 + R_{30}^2 \right] - \frac{1}{8} F_{30}^2 \left[R_{26} + R_{27} + R_{28} + R_{29} + R_{30} + R_{30} + R_{30} + R_{30} + R_{30} \right]^2 \right\}^{\frac{1}{2}}$$

APPENDIX A4

POUT (7,1) = $F_4 R_1$	POUT (7,2) = $F_3 R_2$	POUT (7,30) = $F_{26} R_{30}$
POUT (8,1) = $F_3 R_1$	" 8 " = $F_2 R_2$	" 8 " = $F_{27} R_{30}$
POUT (9,1) = $F_2 R_1$	9 = $F_1 R_2$	9 = $F_{28} R_{30}$
POUT (10,1) = $F_1 R_1$	10 = $F_1 R_2$	10 = $F_{29} R_{30}$
POUT (11,1) = $F_1 R_1$	11 = $F_2 R_2$ ...	11 = $F_{30} R_{30}$
POUT (12,1) = $F_2 R_1$	12 = $F_3 R_2$	12 = $F_{30} R_{30}$
POUT (13,1) = $F_3 R_1$	13 = $F_4 R_2$	13 = $F_{30} R_{30}$
POUT (14,1) = $F_4 R_1$	14 = $F_5 R_2$	14 = $F_{30} R_{30}$
POUT (15,1) = $F_5 R_1$	15 = $F_6 R_2$	15 = $F_{30} R_{30}$

REFERENCES CITED

Filz, R. C., L. Katz, P. L. Rothwell, B. Sellers, F. A. Hansen, and E. Holeman, Observations of 5-45 MeV protons at $L \leq 3$ and $L \geq 7$, Space Res., 14, 321, 1974.

Fischer, H. M., V. W. Auschrat, and G. Wibberenz, Angular distribution and energy spectra of protons of energy $5 \leq E \leq 50$ MeV at the lower edge of the radiation belt in equatorial latitudes, J. Geophys. Res., 82, 537, 1977.

Heckman, H. H. and V. O. Brady, Effective atmospheric losses for 125-MeV protons in South Atlantic anomaly, J. Geophys. Res., 71, 2791, 1966.

Holeman, E., Solar Cycle Variations of Geomagnetically Trapped Radiation, Final Report, Contract No. F19628-69-C-0139, 1973.

Shea, M. A., D. F. Smart, and H. Carmichael, Summary of Rigidities Calculated with the International Geomagnetic Reference Field for Various Epochs, ERP No. 561, AFGL-TR-76-0115, 1976.

Shea, M. A., D. F. Smart, and K. G. McCracken, A Study of Vertically Incident Cosmic-ray Trajectories Using Sixth-degree Simulations of the Geomagnetic Field, ERP No. 141, AFCRL-65-705, 1965.

Sugiura, M., Equatorial current sheet in the magnetosphere, J. Geophys. Res., 77, 6093, 1972.

TABLE I

OBSERVED ENERGY RANGES FOR PROTON FLUXES

Bin	Energy Range	Midpoint	Width(ΔE)	Rigidity
1	5-7 MeV	6 MeV	2.0 MeV	.104 GV
2	7-12.2	9	5.2	.128
3	12.20-18.2	15	6.0	.167
4	18.20-28	23	9.8	.208
5	28-45	36	17.0	.261

TABLE IIa

L = 1.15, B(H_{mm} ~ 300 km) = .214G

B	θ_{LC}	B_s	θ_s	$\Delta\theta$	ΔB
.185G	68°	.697G	31°	37°	.482G
.195	73°	.735	31°	42°	.521
.205	79°	.229	71°	8°	.016

TABLE IIb

L = 1.35, B(H_{mm} ~ 300 km) = .220G

B	θ_{LC}	B_s	θ_s	$\Delta\theta$	ΔB
.185G	66°	.697G	31°	35°	.477G
.195	70°	.735	31°	39°	.515
.205	75°	.292	57°	18°	.072
.215	81°	.254	67°	14°	.034
.225	90°	.242	75°	15°	.017

TABLE IIc

L = 1.55, B(H_{mm} ~ 300 km) = .234G

B	θ_{LC}	B_s	θ_s	$\Delta\theta$	ΔB
.195G	66°	.735G	31°	35°	.501G
.205	70°	.623	35°	35°	.389
.215	74°	.378	49°	25°	.144
.225	79°	.274	67°	12°	.040

TABLE IId

$L = 2.05$, $B(H_{mm} \sim 300 \text{ km}) = .255$ G

B	θ_{LC}	B_s	θ_s	$\Delta\theta$	ΔB
.215G	66°	.306G	57°	9°	.051G
.225	70°	.336	55°	15°	.081
.235	74°	.307	61°	13°	.052
.245	78°	.289	67°	11°	.034

TABLE IIe

$L = 2.75$, $B(H_{mm} \sim 300 \text{ km}) = .285$G

B	θ_{LC}	B_s	θ_s	$\Delta\theta$	ΔB
.245G	68°	.365G	55°	13°	.080G
.255	71°	.363	57°	14°	.078
.265	76°	.346	61°	15°	.061
.275	80°	.346	63°	17°	.061
.285	90°	.336	67°	23°	.051

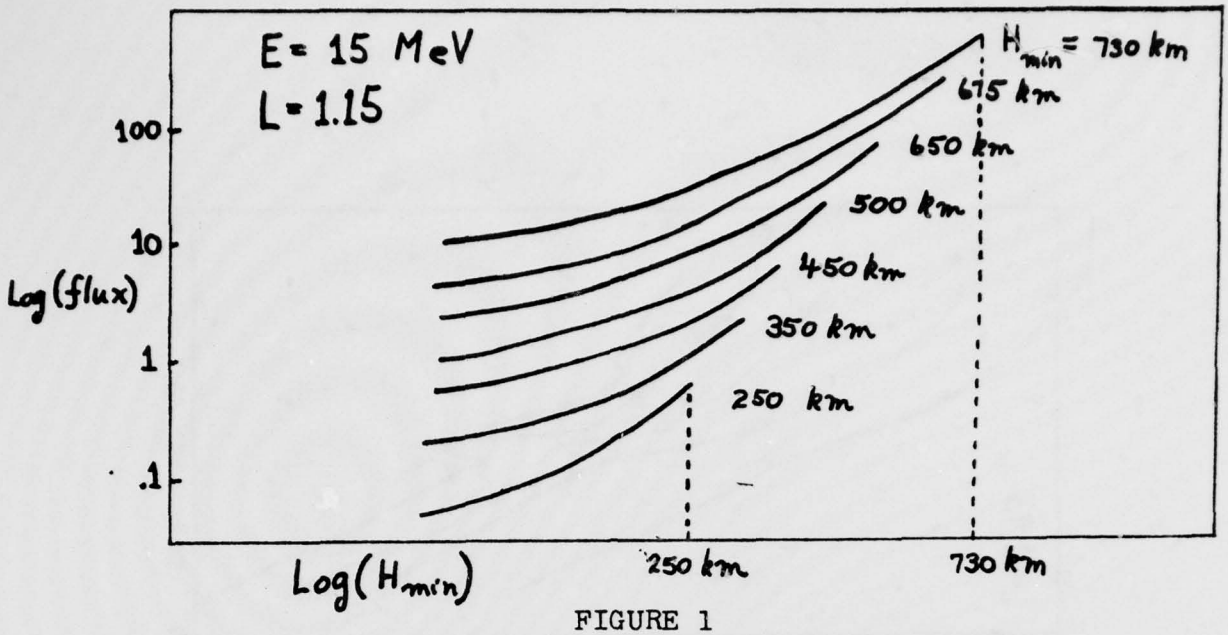


FIGURE 1

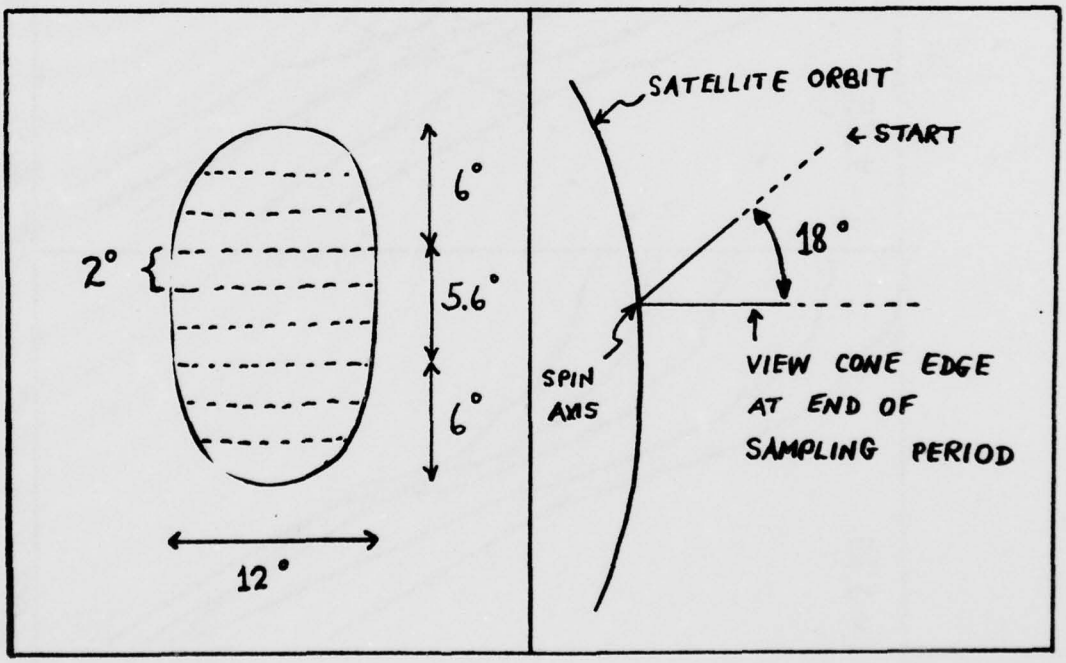


FIGURE 2

FIGURE 3

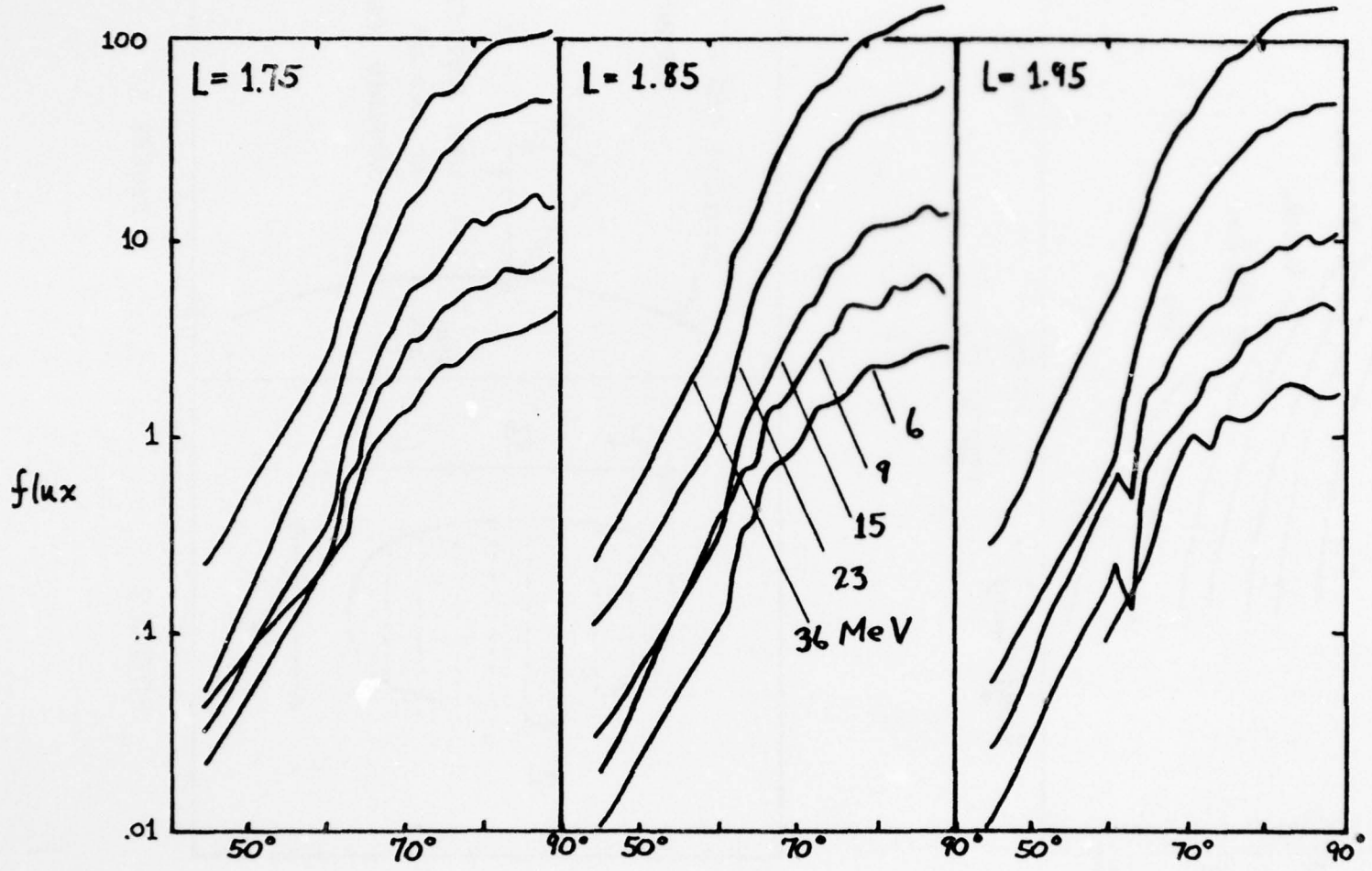


FIGURE 4

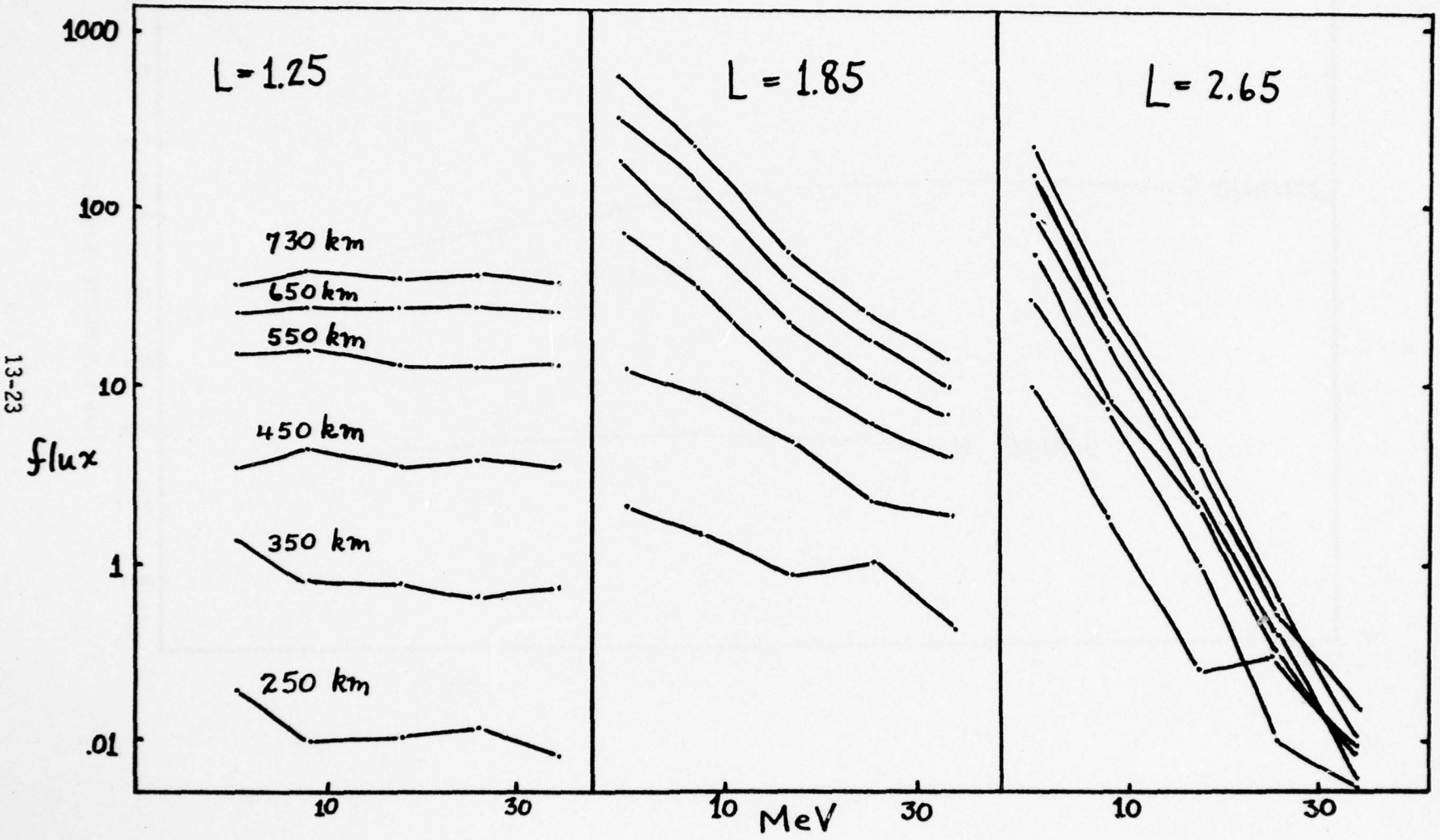

L = 1.25
L = 1.85
L = 2.65
1000
100
10
1
.01
flux
730 km
650 km
550 km
450 km
350 km
250 km
10
30
MeV

FIGURE 5

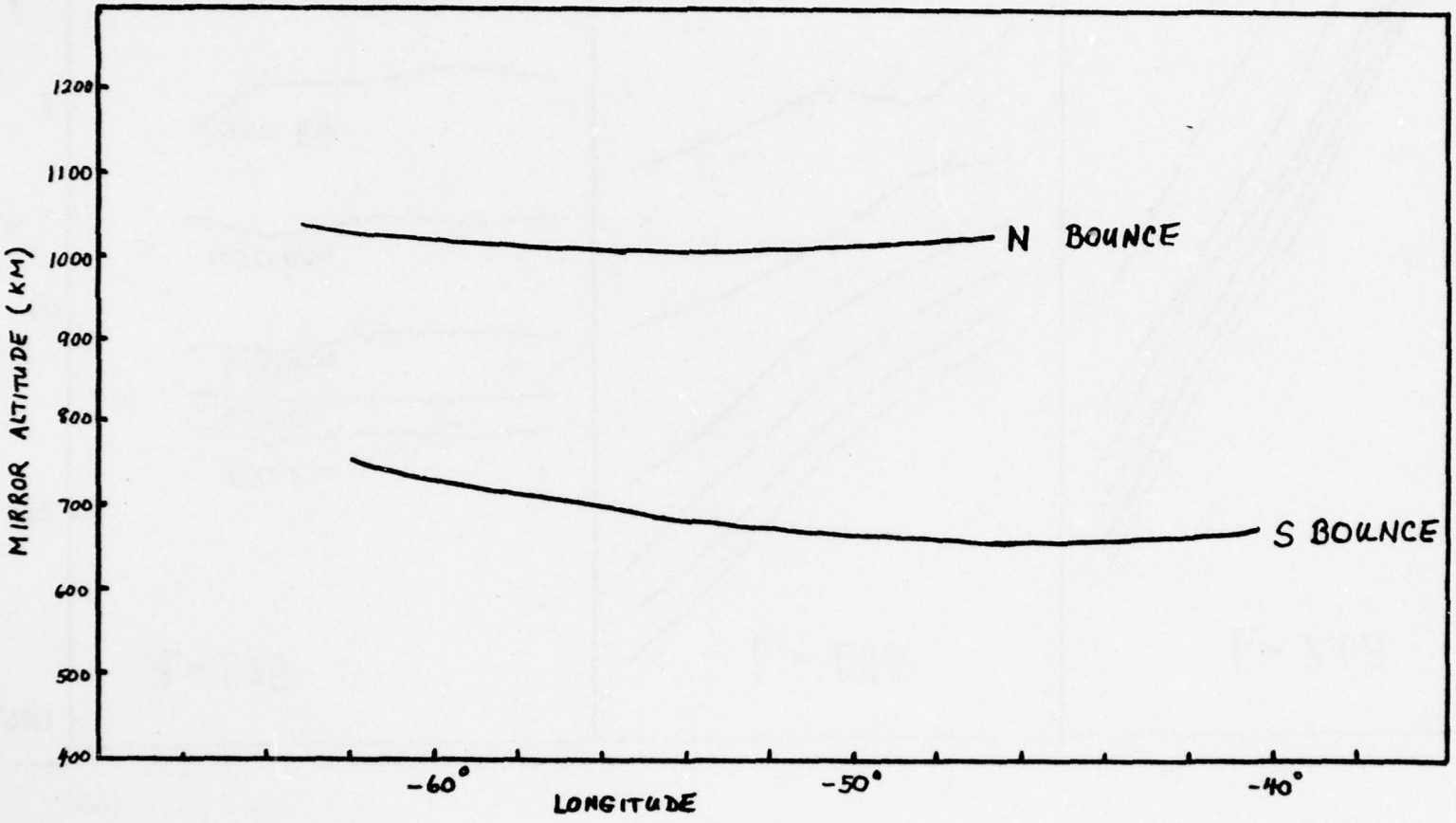

1200
1100
1000
900
800
700
600
500
400
MIRROR ALTITUDE (KM)
N BOUNCE
S BOUNCE
-60°
-50°
-40°
LONGITUDE

FIGURE 6

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

A STUDY OF THE PARTICLE SPECTRUM OF THE CONTINUOUS AURORA AND ASSOCIATED E-REGION IONIZATION BEFORE AND AFTER A MAGNETIC SUBSTORM

Prepared by:	James R. Sharber, Ph D.
Academic Rank:	Associate Professor
Department and University:	Department Physics and Space Sciences Florida Institute of Technology
Assignment:	
Air Force Base	Hanscom AFB
Laboratory	Air Force Geophysics Laboratory
Division	Space Physics Division
Branch	Ionospheric Dynamics
USAF Research Colleague:	James A. Whalen, Ph D.
Date:	August 26, 1977
Contract No.:	F44620-75-C-0031

ABSTRACT

An examination is made of the spectra of electrons and protons responsible for the continuous aurora as measured by the ISIS 2 soft particle spectrometer as the satellite crossed the flight path of the AFGL Airborne Observatory on December 9, 1971. Ionospheric sounder measurements made from the aircraft provide the means of relating the shape and energy content of the particle spectra with the virtual height and critical frequency of the auroral E ionization layer. The electron spectrum of the continuous aurora may be represented by a Maxwellian distribution of energies between a few hundred ev and $\sim$ 10 kev. The characteristic energy decreases with distance from the center of the continuous band, increases after the onset of the substorm, and provides a qualitative indication of the E-layer height. The peak in the curve of energy flux within each accumulation sample of the ISIS 2 electron spectrum provides a quantitative indicator of the E-layer height. The ratio of $(foE)^4$ before and after the substorm is approximately the same as the ratio of electron energy deposited within the E-region, before and after the substorm. Protons contributed less than 10% to the energy deposited in the E region. Their average energy was four to five times the average electron energy.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude at having been provided the opportunity to work as a participant in the 1977 USAF-ASEE Summer Faculty Research Program sponsored by the Air Force Office of Scientific Research (Contract No. F44620-75-C-0031) and conducted by Auburn University. I especially wish to thank Mr. Fred O'Brien, Program Director, for his help and guidance throughout the program.

I wish to thank my research colleague Dr. J.A. Whalen and the other members of the Ionospheric Dynamics Branch of AFGL, in particular, Dr. E.J. Weber, Ms. R.A. Wagner and Dr. J.R. Jasperse and Lt Col A.L. Snyder for their very useful discussions and assistance during this study. I wish also to thank Mr. Bob Boudreau for programming the ISIS spectral plots and Mr. Bill Wong for technical assistance. Much of the ISIS data was provided by U.T. Dallas; my thanks are extended to Dr. J.D. Winningham, Dr. Dave Klumpar and Mrs. Carolyn Stuart for their fast response to my requests. The work was supported in part by contract AFGL-TR-77-0047.

INTRODUCTION

As a result of several flights of the AFGL Airborne Ionospheric Observatory (Whalen et al., 1971; Weber et al., 1977; Whalen et al., 1977) and polar satellite observations (Lui et al., 1977) it has become clearly established that a band of continuous auroral emission exists equatorward of and sometimes overlaps the discrete auroral oval. Further, this band of diffuse and sometimes subvisual auroral emission corresponds very closely with 'auroral-E,' a relatively thick E-region ionization caused primarily by incoming auroral particles (Whalen et al., 1971; Whalen et al., 1977).

In the present report an examination is made of the spectra of electrons and protons responsible for the continuous aurora as measured by the ISIS 2 soft particle spectrometer as the satellite crossed the flight path of the AFGL Airborne Observatory on December 9, 1971. Ionospheric sounder measurements made from the aircraft provide the means of relating the shape and energy content of the particle spectra with the virtual height and critical frequency of the auroral E ionization layer.

A number of observations made during this midnight sector flight have been reported as three-paper case study of auroral activity (Whalen et al., 1977), electrojet and precipitating particle characteristics (Weber et al., 1977), and ionospheric F-region phenomena (Pike et al., 1977). As a result of this case study the documentation of geophysical activity during the flight is quite thorough, and since the period includes two substorm sequences, it is felt that an excellent opportunity is provided for the present study. In our investigation we concentrate on times before and after the second (and more classic) substorm. This enables a comparison

of the energy spectra and ionospheric properties measured during the quiet, pre-substorm period with those characteristic of the expansive phase.

OBJECTIVES

The objectives of the study are:

1. To describe the electron and proton energy spectra which produce the continuous aurora observed before and after the 0847 UT substorm of 09 December 1971.

2. To determine whether a spectral parameter such as total energy, peak energy, average energy, or energy within a specified range may be correlated with the virtual height and critical frequency of auroral-E ionization.

AIRCRAFT IONOSPHERIC MEASUREMENTS

The aircraft flew from Goose Bay, Labrador to Eielson AFB, along the trajectory shown as the heavy line in Figure 1 (from Whalen et al., 1977). Its course and speed were adjusted so that the aircraft remained in the 22-00 MLT sector and intersected the DMSP (solid lines) and ISIS 2 (dotted lines) trajectories as shown. In the figure each satellite pass is labeled by the universal time at which it crossed the aircraft flight path.

During the flight, ionospheric soundings were made once each minute with the Granger Model 3905-1 sounder. Each measurement required 20 sec to sweep the frequency range 2 to 8 MHz. The sounder data which pertain to the E region are shown in the graphs of Figure 2 (compliments of R.A. Wagner) and are labeled foE (MHz) and h'E (km). The quantity foE is the plasma critical frequency; its fourth power, also plotted in Figure 2, is approximately proportional to the total energy deposited in the E region which produces ionization

(Omholt, 1955; Whalen et al., 1971). The quantity h'E is the virtual height of the ionization layer, or the height above the ground of the lowest ledge of E region ionization. It is the parameter h'E which gives some indication of the hardness of the incoming particle spectrum since high-energy particles are expected to penetrate further into the atmosphere. For example the increase in h'E between 0720 and 0820 suggests a corresponding decrease in the energy of the particles producing the layer above the aircraft.

We note from Figure 2 that generally, as h'E decreases, foE (and $(foE)^4$) increases and vice versa. This suggests that an increase in energy deposited into the E-region results from an increase in the energy of the precipitated particles.

Another feature of Figure 2 is the change in the E-region parameters before and after the onset of the substorm at 0847. (Figure 3 shows both the AE index and the College magnetogram.) The virtual height measured by the aircraft sounder is generally higher before the substorm onset than after. Correspondingly, foE gradually diminishes before the substorm to a value of ~ 2.0 MHz at 0815, then reaches 3.8 MHz at about 1000 UT. Between 0815 and 0920 UT the aircraft was poleward of the continuous aurora resulting in no detection of auroral E during this time.

It is to be noted that as the sounder data of Figure 2 were being taken, the aircraft underwent changes in corrected geomagnetic latitude in accordance with Figure 1. The variations in foE and h'E shown in Figure 2 can therefore arise from any (or from combinations) of the following: substorm phase, motion of the region of particle precipitations, temporal or spatial change in the precipitating particle characteristics, motion of the aircraft relative to precipitation region. Some of these phneomena have been described in the study of Whalen et al. (1977) and details of the complex

motion of the precipitation region are currently under study (J.A. Whalen, private communication). These phenomena must be considered as we compare the data of Figure 2 with spectra measured by the ISIS-2 satellite.

ISIS 2 SOFT PARTICLE SPECTROMETER

The particle observations were made from ISIS 2, a high-inclination polar satellite with a nearly circular orbit of mean altitude 1400 km. The soft particle spectrometer is similar to the ISIS 1 spectrometer (Heikkila et al., 1970) and uses a divergent plate, electrostatic deflection system to discriminate energies of electrons and positive ions (hereafter called protons) in the nominal energy range 5 ev to 13 kev. The instrument is composed of two completely separate deflection systems, one positioned atop the other, each with its own collimation system, deflection plates, and particle sensors (discrete dymode multipliers). In the normal mode of operation electrons are analyzed by the top system and protons by the bottom system. The energy resolution ($\Delta E/Eo$) of each deflection system is constant as a function of energy and is 0.247 for the top (electron) analyzer and 0.355 for the bottom (proton) analyzer. The geometric factors are 3.95×10^{-4} cm^2 ster for the electron analyzer and 1.27×10^{-3} cm^2 ster for the proton analyzer.

A spectrum is measured by sampling the detector outputs successively as the potential difference between the deflection plates decays exponentially through values which correspon to energies of 13.2 kev to 5.5 ev for electrons and 14.7 kev to 6.2 ev for protons. The center energies of the samples, the sample bandwidths, and the detection efficiencies of the multipliers are listed in Table 1. A complete spectral measurement of electrons and protons is made each second.

TABLE 1. ISIS 2 SOFT PARTICLE SPECTROMETER CHARACTERISTICS

	ELECTRON DETECTOR				PROTON DETECTOR		
Sample	Center Energy Eo(eV)	Bandwidth ΔE (ev)	Efficiency η	Sample	Center Energy Eo(eV)	Bandwidth ΔE (ev)	Efficiency η
1	13150	3248	.230	1	14675	5210	.8
2	12620	3117	.235	2	14084	5000	.8
3	9805	2422	.270	3	10942	3884	.8
4	7919	1956	.295	4	8837	3137	.8
5	6395	1580	.320	5	7137	2534	.8
6	5165	1276	.350	6	5764	2046	.8
7	4171	1030	.375	7	4655	1653	.8
8	3369	832	.400	8	3759	1334	.8
9	2721	672	.430	9	3036	1078	.8
10	2197	543	.460	10	2452	870	.8
11	1774	438	.495	11	1980	703	.8
12	1433	354	.535	12	1599	568	.8
13	1157	286	.575	13	1292	459	.8
14	935	231	.620	14	1043	370	.8
15	755	186	.665	15	842	299	.8
16	610	151	.705	16	680	241	.8
17	492	122	.735	17	549	195	.8
18	398	98.3	.760	18	444	158	.8
19	321	79.3	.778	19	358	127	.8
20	259	64.0	.790	20	289	103	.8
21	209	51.6	.800	21	234	83.1	.8
22	169	41.7	.800	22	189	67.1	.8
23	137	33.8	.800	23	152	54.	.8
24	110	27.2	.790	24	123	43.7	.8
25	89.1	22.0	.782	25	99.4	35.3	.8
26	72.0	17.8	.772	26	80.3	28.5	.8
27	58.1	14.4	.764	27	64.9	23.0	.8
28	46.9	11.6	.755	28	52.4	18.6	.8
29	37.9	9.36	.748	29	42.3	15.0	.8
30	30.6	7.56	.738	30	34.2	12.1	.8
31	24.7	6.10	.730	31	27.6	9.8	.8
32	20.0	4.94	.721	32	22.3	7.92	.8
33	16.1	3.98	.715	33	18.0	6.39	.8
34	13.0	3.21	.710	34	14.5	5.15	.8
35	10.5	2.59	.705	35	11.7	4.15	.8
36	8.5	2.10	.700	36	9.5	3.37	.8
37	6.9	1.70	.700	37	7.7	2.73	.8
38	5.5	1.36	.700	38	6.2	2.20	.8

PARTICLE OBSERVATIONS

Of the ISIS 2 passes shown in Figure 1, our consideration will be restricted primarily to those labeled 0610, 0802, and 0954 UT. We have made use of some data from the 0418 pass, but on this pass the satellite was only tracked as it crossed the most equatorward portion of the continuous band. In this section the general characteristics of the electron and proton precipitation are presented in the form of ISIS 2 spectrograms. Details of the energy spectra will be discussed in the following section. A discussion of the general aspects of the particle precipitation during these times has been reported by Weber et al. (1977).

The time which includes the 0610 satellite pass is a quiet stable period between the two substorms shown in Figure 3. At 0610 UT the continuous aurora extended from 67° to 72.6° corrected geomagnetic latitude as is indicated by Figure 4, which provides a universal time history of the region of the continuous aurora as determined by all-sky camera pictures, sounder measurements, DMSP images, and ISIS 2 electron data (Whalen et al., 1977).

The ISIS 2 electron and proton data for the complete pass over the continuous auroral region are shown as energy-time spectrograms of Figure 5. In this presentation each vertical line in the top portion of the graphs represents a differential energy spectrum (ergs/cm^{2}ster sec). Intensity is displayed by gray-scale shading with high intensities represented as lighter shades of gray. The logarithm of energy is the ordinate and universal time, magnetic local time, and invariant latitude are the abscissas. Also shown in the two graphs below the spectrograms are the number and energy fluxes obtained by integration over the range of each spectrum.

Electron spectra show very little structure between invariant latitudes of 73° and 66°. This is characteristic of the continuous aurora (Meng, 1976; Deehr et al., 1976; Lui et al., 1977) and is confirmed in this case by the lack of any discrete arcs observed at this time by the airborne all-sky camera (Whalen et al., 1977).

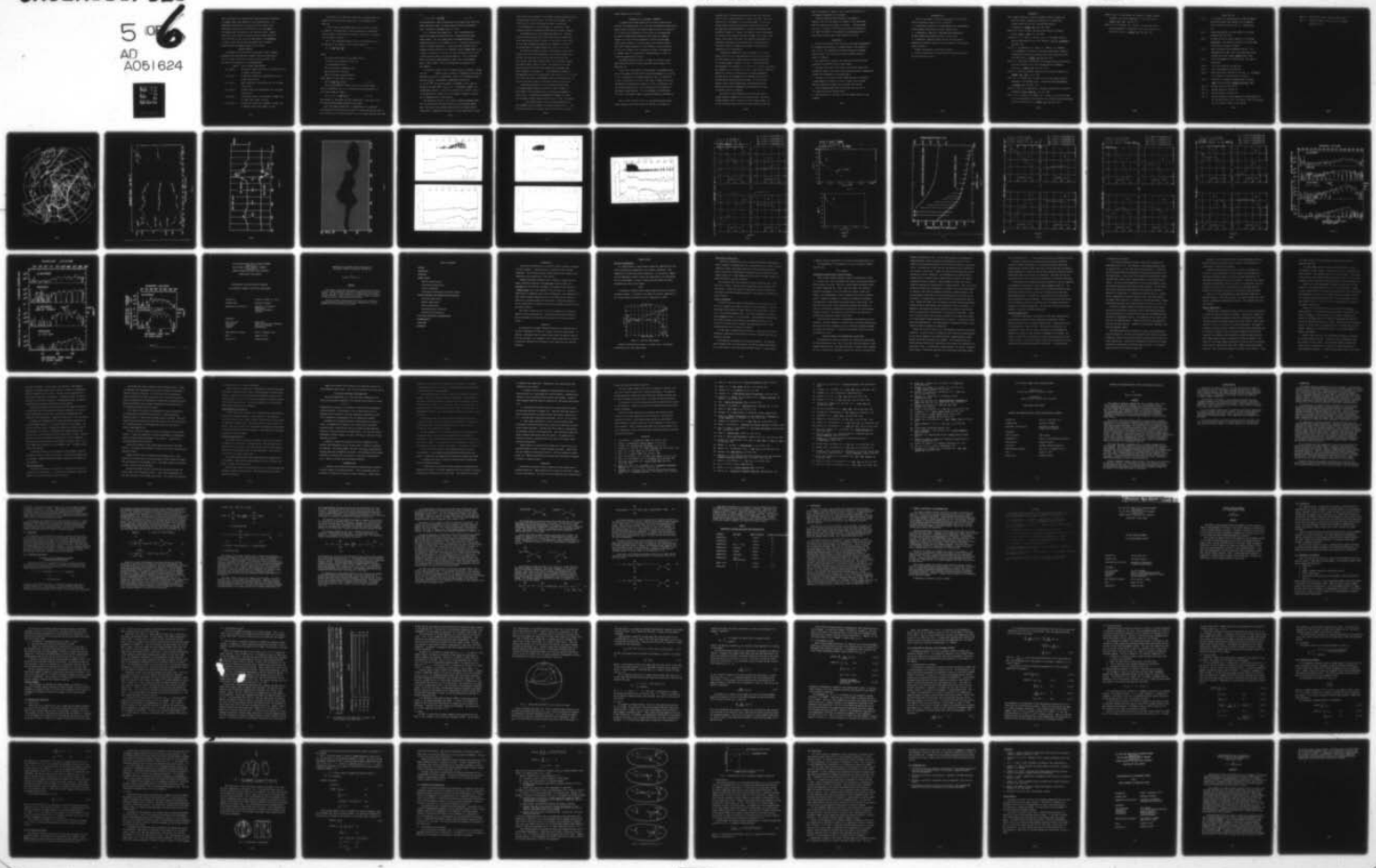
Protons are observed between latitudes of 78.5° and 66°. The energy carried by the protons is low, ≤10% of that of the electrons. The proton average energy increases with decreasing latitude down to about 68° where the broad spectral peak appears near the upper end of the analyzer energy range.

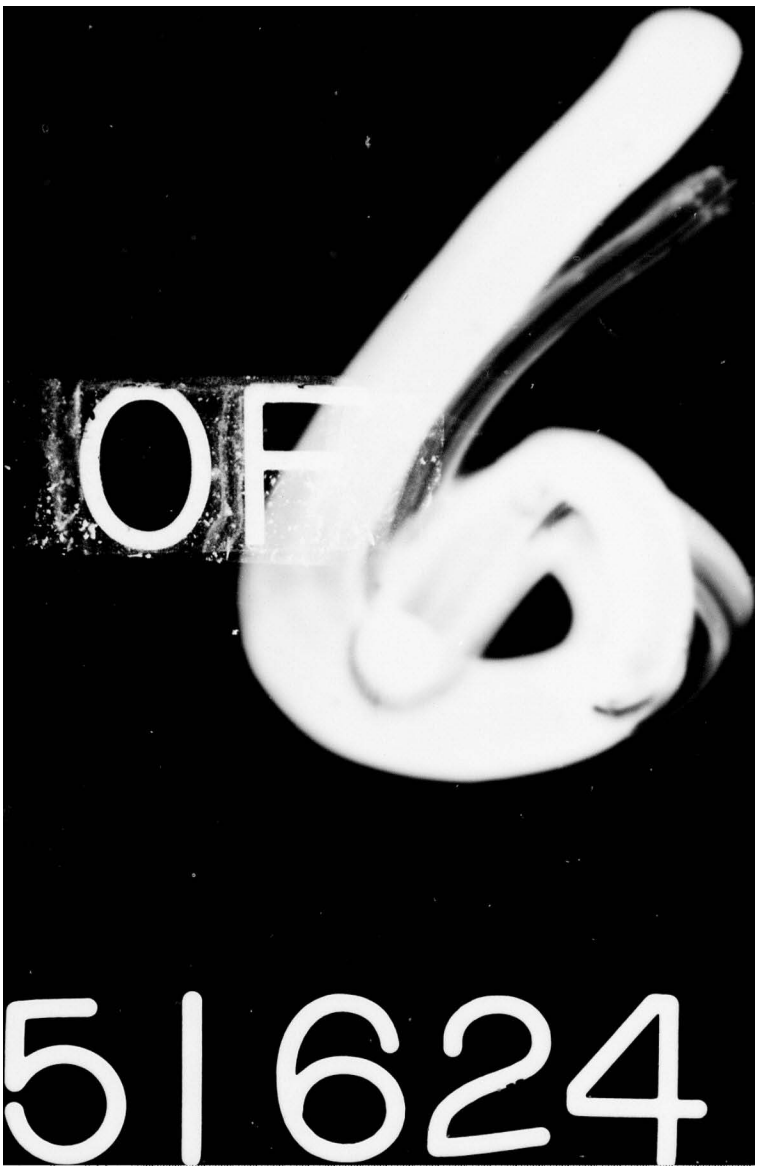
The ISIS pass which crossed the aircraft flight path at 0802 is shown in Figure 6. The electrons associated with continuous and discrete auroras are observed between 71.5° and 66°. It is clear from the peaking in the electron spectrogram that discrete arcs were present as was reported by Whalen et al. (1977)and Weber et al. (1977). It is also apparent that the electron spectra are somewhat softer than those of the 0610 pass (Figure 5). This observation is in qualitive agreement with the increase in h'E of Figure 2. Such observations will be discussed more quantitatively as the spectra are presented in the next section. The auroral protons are present over about the same range of latitudes as the electrons. As in the 0607 pass their spectra peak broadly in the kev range with the average energy increasing with decreasing latitude. The particles apparent in the spectrograms equatorward of about 65° are outer belt radiation.

Although the proton data were not available for the 0956 UT ISIS pass, the electron spectrogram is shown in Figure 7. The precipitation region extends from about 70° to 64° invariant latitude. As can be

AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

5 OF 6
AD
A051624




OF
51624

seen in the figure, the intensity and energy carried by the electrons is somewhat higher than observed on the preceeding orbits. Two enhancements characteristic of auroral arcs are present on the spectrogram at latitudes of 68.8° and 68.0°, but they appear to be immersed within the spectra of the continuous aurora. Figure 4 shows that at this time the continuous aurora extended from 64.3° to 69° CGL which is born out by the spectrogram. The magnetogram of Figure 3 shows that the satellite passed over the auroral activity during the expansive phase of the substorm.

PARTICLE SPECTRA

In carrying out the objectives of the present study, a computer program has been developed which computes, prints, and plots any or all of six forms of the ISIS 2 electron or proton spectrum. The plot options are presented in the following table.

TABLE 2. ISIS 2 SPECTRAL PLOT OPTIONS

Plot Option 1	Differential number flux (elec or prot/cm^2ster sec ev) vs energy, log/log scale
Plot Option 2	Differential energy flux (ergs/cm^2ster sec ev) vs energy, log/log scale
Plot Option 3	Sample energy flux (ergs/cm^2ster sec ev) vs energy, log/log scale
Plot Option 4	Integral energy flux (ergs/cm^2ster sec) vs energy, log/log scale
Plot Option 5	(Differential number flux ÷ energy) vs energy, log vs linear scale, Range: 0-14 kev
Plot Option 6	(Differential number flux ÷ energy) vs energy, log vs expanded linear scale, Range: 0-1 kev

The program also provides the versatility of summing number and energy fluxes over arbitrary energy bands of the measured ISIS 2 spectrum.

A set of spectra representing each option of Table 2 is shown in Figure 8. The spectra are of electrons measured at 06/10/48 UT (hour 6, minute 10, sec 48) and are labeled 1 through 6 to coincide with the options of Table 2.

Plot 1 the differential number flux, J_N (elec/cm^{2}ster sec ev) is obtained for each sample by the following expression

$$J_N = \frac{C}{G \cdot \Delta t \cdot \eta \cdot \Delta E} \qquad \text{----- (1)}$$

where

E_o is the center energy of each sample (Table 1)

C is the count accumulated in time Δt

G is the geometric factor

Δt is the sample accumulation time

η is the detector efficiency

ΔE is the energy bandwidth (FWH M)

Since $\Delta E/Eo = R$, the energy resolution,

$$J_N = \frac{C}{G \cdot \Delta t \cdot \eta \cdot R \cdot Eo} \qquad \text{----- (2)}$$

Plot 2 the differential energy flux J_N, (or spectral energy density (ergs/cm^{2}ster sec ev) is obtained by multiplying equation (2) by the mean energy per sample

$$J_E = J_N \cdot E_o = \frac{C}{G \cdot \Delta t \cdot \eta \cdot R} \qquad \text{----- (3)}$$

This quantity is very nearly directly proportional to the count C since only the efficiency varies (slightly) with energy.

Plot 3 the sample energy flux, E (ergs/cm^{2}ster sec), represents the energy contained within a given sample of energy width ΔE (See Table 1) and is obtained by multiplying equation (3) by the energy bandwidth ΔE ($=RE_o$).

or

$$E = J_E \cdot \Delta E = \frac{C \quad Eo}{G \ \Delta t \cdot \eta} \qquad \text{----- (4)}$$

This representation should be appropriate for E-region study since each point specifies an amount of energy deposited within an atmospheric layer. The peak in the graph occurs at 2.2 kev.

Plot 4 displays the integral flux $> Eo$, (ergs/cm^{2}ster sec) obtained by summing the values from Plot 3 starting with the high-energy channel. The total energy under the measured spectrum is indicated by the value of the graph at 10 ev. This plot is particularly suited to E-region application. A practical upper altitude for the E-region is about 150-160 km; the ionization curves of Rees (1964) for an isotropic distribution of electrons at 1-kev energy place the altitude of peak ionization rate at 156 km (See Figure 9). Plot 4 shows that 83% of the electron energy measured by ISIS 2 lies in the channels above ~ 1000 ev range. It is this energy which is deposited into the E-region.

In the plots labeled 5 and 6 the quantity J_N/E_o is plotted vs energy with the ordinate, J_N/E_o, plotted on a logarithmic scale and the abscissa E_o plotted on a linear scale. Thus if the differential number spectrum J_N can be represented as a Maxwellian distribution $J_N = N_o \, E \, e^{-E/\alpha}$, the resulting graph is a straight line of slope m related to α (=kT) by $\alpha = (\log_{10} e)/m$. In the present example, the characteristic energy is 740 ev with an No of $5.6 \times 10^2/cm^3$. Plot 6 is the same as plot 5 except the 0-1 kev range is expanded to analyze the low energy porition of the spectrum.

The spectrum of 06/10/48 UT is typical of spectra measured within the loss cone but near the peak intensity of the continuous auroral band. That part of the spectrum deposited into the E-region may be represented as a Maxwellian distribution, but the characteristic energy

(740 ev) does not correspond to the height of maximum production rate of ionization which was much lower. The value of h'E (Figure 2) was ~125 km at this time, which places the peak ion density at 133-135 km, applying the curve of Figure 9. The values of total energy and number flux measured at 06/10/48 UT are shown in Figure 8.

The highest value of total energy and number flux measured within the loss cone during the ISIS 2 0610 UT pass occurred about 22 sec earlier and is shown in Figure 10 (only plots 1-4 are shown). At this time the Maxwellian characteristic energy is 1.09 kev and the energy peak of the sample energy flux of Plot 3 is 4.2 kev.

The spectrum presented in Figure 11 was measured at 08/03/04 UT during a time when foE was relatively low and virtual heights were ~150 km. This spectrum is less intense than those during the 0610 UT pass; both number flux and energy flux at E-region energies are lower by about a factor of three. The spectrum is also much softer with only 30% of the total energy within the range above 935 ev (channel 14, see Table 1.). The characteristic energy of the Maxwellian portion of the spectrum is 490 ev with $N_o = 2.2 \times 10^2$ /cm^3. None of the plots has a very prominent peak at E-region energies, but in plot 3 the curve contains a broad peak between 300 and 500 ev.

A spectrum representative of the continuous aurora during the substorm expansive phase at 09/54/57 is shown in Figure 12. The presence of high energy particles is particularly apparent in plots 3 and 4. The Maxwellian characteristic energy is 1.111kev with the peak of sample energy flux (plot 3) of 2.4kev. The energy fluxes exceed those in the hour before substorm onset (the 08/03/04 UT pass of Figure 11) by about a factor of four . The aircraft measurements of h'E give values of 115-117 km, the lowest

heights measured by the aircraft.

COMPARISON WITH IONOSPHERIC PARAMETERS

In comparing the virtual heights with the ISIS 2 spectra before and after the substorm onset it is clear that the Maxwellian characteristic energy and also the peak in the sample energy flux (plot 3 of Figures 8, 10, 11, and 12) are indicators of the altitude of the E-layer peak. The peak of the Maxwellian distribution just before the substorm (Figure 11) is 450 ev which corresponds to 198 km using the Rees (1964) curve of Figure 9. The lower edge of the ionization profile which more closely represents h'E, may be taken to be the 10% point on the curve of Figure 9 and corresponds to 167 km. The edge would be somewhat lower if the spectrum is not monoenergetic; a value of 150 km was measured by the sounder.

The sample energy flux (plot 3 of Figure 11) contains a broad peak between 300 and 500 ev which agrees well with the peak of the Maxwellian fit.

At the time of the expansive substorm phase, the parameter which most closely agrees with the aircraft sounder measurement of 115-117 km is the 2.4 kev peak of the sample energy flux (plot 3 of Figure 12). In fact, the 10% value at 3.2 kev on the curve of Figure 9 is 122 km. The Maxwellian characteristic energy of 1.11 kev is clearly too low, but results because the fit did not include the two data points of the highest energy channels. It is not uncommon in the ISIS data that flux in the highest channels exceeds the Maxwellian distribution.

While a clear difference exists in the 08/03/04 spectrum (just before substorm onset) and the 09/54/57 spectrum (just after the

substorm onset), the spectrum measured well before the substorm at 06/10/26 shows a marked similarity to that of 09/54/57. Thus there appears to be a discrepancy in the expected height of the auroral E-layer based on spectral measurements during the 0610 UT ISIS pass and the aircraft measurements of virtual height between 0608 and 0612 UT (Figure 2). However, the Churchill sounder measurements at the time of the satellite pass show a height of less than 110 km. Figure 1 shows that at the time of the satellite measurement, the longitude of the satellite was between that of the aircraft and Churchill; therefore a longitudinal gradient in height might be responsible for the discrepancy.

A qualitative comparison is made in Figures 13, 14, and 15 between $(foE)^4$ and the energy deposited into the auroral E-region. The ISIS energy fluxes of these figures are taken from the data of the spectrograms of Figures 5, 6, and 7. The electrons have been integrated over channels 1 to 14 (935 ev to 13.2 kev), the protons over channels 1 to 9 (3.0 kev to 14.7 kev). The values of $(foE)^4$ are plotted as a function of the aircraft's corrected geomagnetic latitude and do not correspond to the universal times shown for the ISIS data at the bottom of each figure. It is clear from Figures 13 and 15 that an approximate proportionality exists between the two kinds of measurements although they were made at different times and locations.

Also shown in these figures are the electron and proton average energies, averaged over each spectral sweep. The proton average energy is about 4 times the electron average energy. As we have previously noted, the protons usually contribute less than

10% of the energy, but Figures 13 and 14 show that almost all of this energy is deposited in the E-region.

Using the numerical data displayed on the graphs of Figures 11 and 12 we can compare the change in $(foE)^4$ before and after the substorm to the change in the total electron energy flux, E_T . The ratio of $(foE)^4$ before and after the substorm is $46.2\ (MHz)^4/208\ (MHz)^4 = 0.22$, whereas the corresponding ratio of the electron energy fluxes is 0.27 erg/1.11 erg = 0.24.

CONCLUSIONS

1. The electron spectrum of the continuous aurora may be represented as a Maxwellian distribution of energies between a few hundred ev and ~10 kev. The characteristic energy decreases with distance from the center of the continuous band and increases after the onset of a substorm.
2. The characteristic energy of the Maxwellian distribution gives a qualitative indication of the E-layer height.
3. The peak in the curve of the energy flux contained within each accumulation sample of the ISIS electron spectrum provides a quantitative indicator of the height of the auroral E-layer.
4. The ratio of the quantity $(foE)^4$ before and after the substorm is approximately the same as the ratio of total electron energy deposited within the E-region before and after the substorm.
5. Proton average energy before the substorm onset was four to five times the electron average energy.
6. Protons contribute less than 10% of the energy deposited in the E-region.

RECOMMENDATIONS

With the development of the plotting capability for the ISIS 2 spectra the present study should be followed by:

1. A simultaneous comparison of plasma sheet particle measurements, airborne sounder measurements and ISIS 2 particle data.
2. A simultaneous comparison of Chatanika Radar measurements, aircraft sounder measurements and ISIS 2 particle data.
3. A theoretical computation which uses the incident particle spectrum as a boundary condition in the calculation of true electron density profiles.
4. Further investigation of the proton component.
5. Further investigation of the acceleration processes associated with the continuous aurora.

REFERENCES

Deehr, Charles Sterling, J. David Winningham, Fumihiko, Yasuhara, and Syun-Ichi Akasofu, Simultaneous observations of discrete and diffuse auroras by the ISIS 2 satellite and airborne instruments, J. Geophys. Res., 81, 5527, 1976.

Eather, R.H. and K.M. Burrows, Excitation and ionization by auroral protons, Aust. J. Phys., 19, 309, 1966.

Heikkila, W.J., J.B. Smith, J. Tarstrup, and J.D. Winningham, The soft particle spectrometer in the ISIS-1 satellite, Rev. Sci. Instruments, 41, 1393, 1970.

Lui, A.T.Y., D. Vankatesan, C.D. Anger, S.-I. Akasofu, W.J. Heikkila, J.D. Winningham, J.R. Burrows, Simultaneous observations of particle precipitations and auroral emissions by the ISIS-2 satellite in the 19-24 MLT sector, J. Geophys. Rev. 82, 2210, 1977.

Meng, C.-I., Simultaneous observations of low-energy electron precipitation and optical auroral arcs in the evening sector by the DMSP-32 satellite, J. Geophys. Res., 81, 2771, 1976.

Omholt, A., The auroral E-layer ionization and the auroral luminosity, J. Atmosph. Terr. Phys., 7, 73, 1955.

Pike, C.P., J.A. Whalen, and J. Buchau, A 12 hour case study of auroral phenomena in the midnight sector: F-layer and 6300 Å measurements, J. Geophys. Res., 82, 3547, 1977.

Rees, M.H., Note on the penetration of energetic electrons into the earth's atmosphere, Planet. Space Sci., 12, 722, 1964.

Weber, E.J., J.A. Whalen, R.A. Wagner, and J. Buchau, A 12 hour case study of auroral phenomena in the midnight sector: Electrojet and precipitating particle characteristics, J. Geophys. Res., 82, 3557, 1977.

Whalen, James A., Jurgen Buchau, and Rosemarie A. Wagner, Airborne ionospheric and optical measurements of noontime aurora, J. Atmosph. Terr. Phys., 33, 661, 1971.

Whalen, J.A., R.A. Wagner, and J. Buchau, A 12 hour case study of auroral phenomena in the midnight sector: Oval, polar cap, and continuous auroras, J. Geophys. Res., 82, 3529, 1977.

FIGURE CAPTIONS

Fig. 1. A CG latitude, CG longitude map of the AFGL Ionospheric Observatory flight track together with the DMSP and ISIS 2 satellite tracks of 9 Dec. 71. Labels are in UT.

Fig. 2. Sounder measurements of virtual height h'E, critical frequency foE, and $(foE)^4$.

Fig. 3. AE index and the horizontal component of the College magnetogram. The numbers in the center are the UT times of the satellite passes of Figure 1.

Fig. 4. Band of continuous aurora observed on 09 Dec. 71 by measurements from DMSP, ISIS 2, and ground and airborne sounders and all-sky cameras (Whalen et al., 1977).

Fig. 5. ISIS 2 spectrogram for the 0610 UT pass. See text for explanation.

Fig. 6. ISIS 2 spectrogram for the 0802 pass.

Fig. 7. ISIS 2 spectrogram for the 0954 pass.

Fig. 8. Electron spectra of 06/10/48 UT of 09 Dec. 71. The numbers 1-6 correspond to the plot options of Table 2.

Fig. 9. Production of ionization by monoenergetic electrons (Rees, 1964) and protons (Eather and Burrows, 1966).

Fig. 10. Electron spectra of 06/10/26 UT.

Fig. 11. Electron spectra of 08/03/10 UT.

Fig. 12. Electron spectra of 09/54/27 UT.

Fig. 13. Electron and proton average energy and E-region energy flux for the 0610 UT ISIS 2 pass. The quantity $(foE)^4$ as determined from the aircraft sounder is also plotted.

Fig. 14. 0802 UT ISIS 2 pass. Otherwise same as Fig. 13.

Fig. 15. 0954 ISIS 2 pass for electrons only. Otherwise same as Fig. 13.

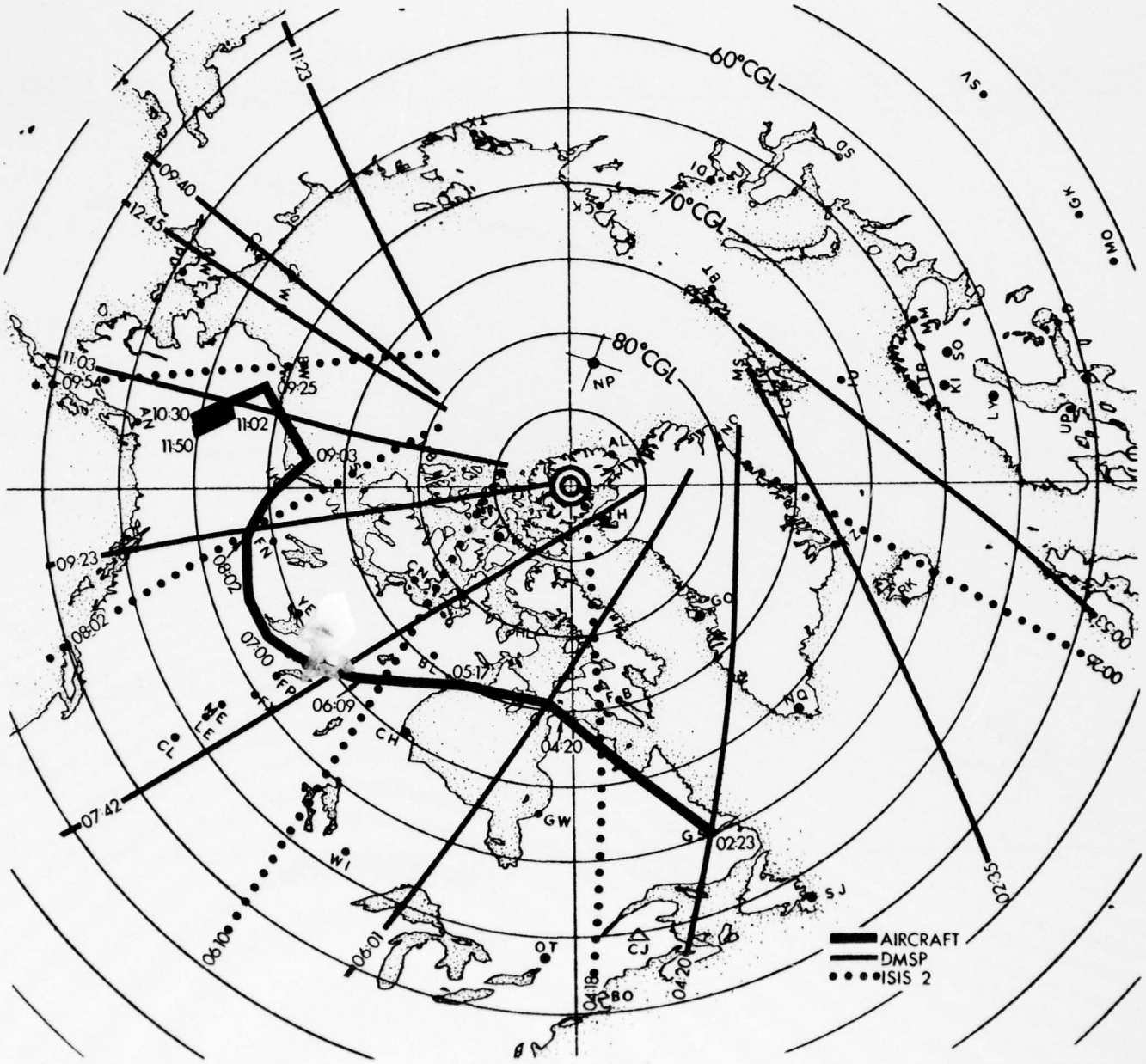

60°CGL
70°CGL
80°CGL
NP
AIRCRAFT
DMSP
ISIS 2
11:23
09:40
12:45
11:03
09:54
10:30
11:50
11:02
09:25
09:03
09:23
08:02
08:02
07:00
05:17
06:09
04:20
07:42
06:10
06:01
02:23
00:53
02:35
00:28

Figure 1

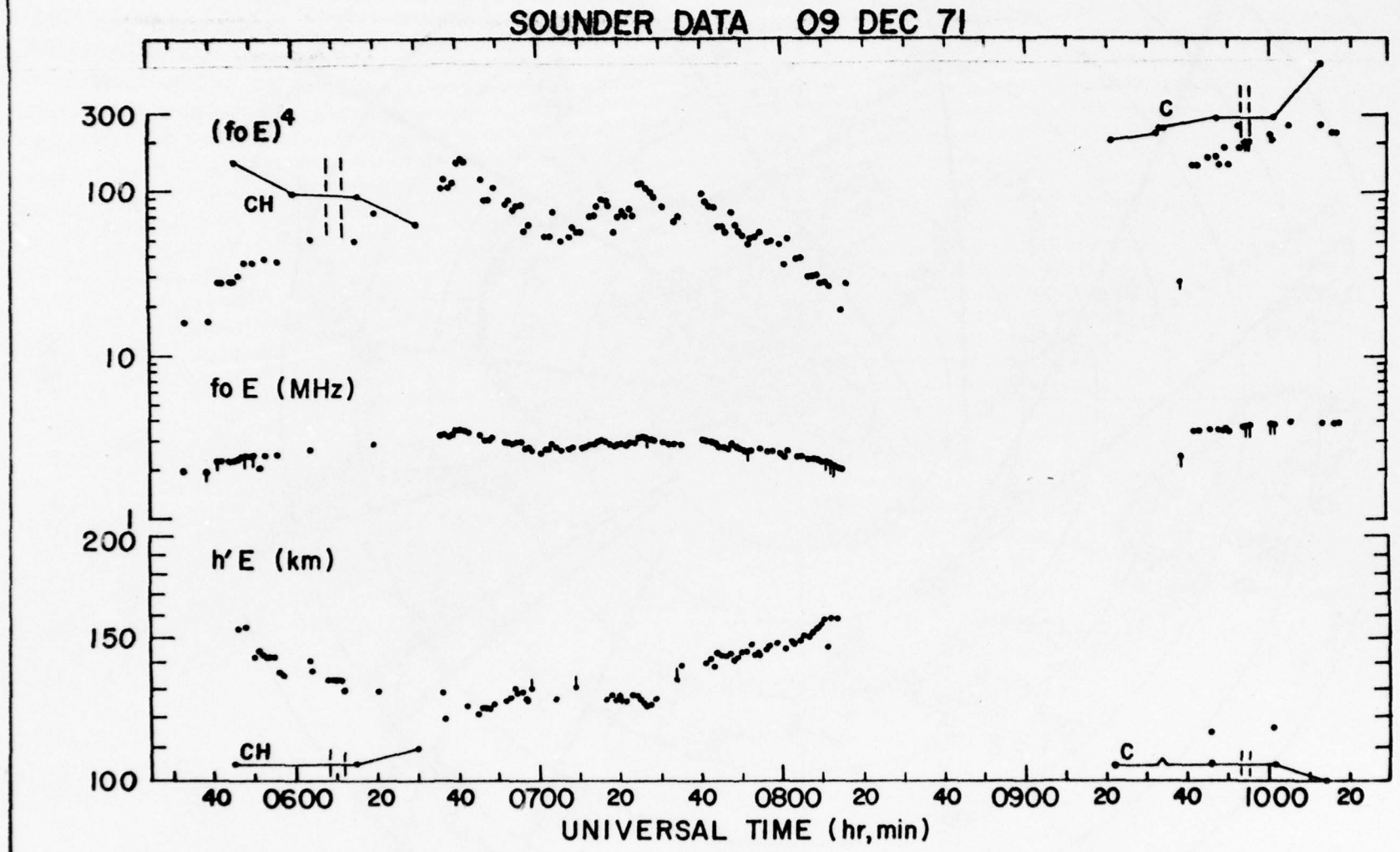

SOUNDER DATA 09 DEC 71
300
100
10
1
(foE)4
CH
C
foE (MHz)
200
150
100
h'E (km)
CH
C
40
0600
20
40
0700
20
40
0800
20
40
0900
20
40
1000
20
UNIVERSAL TIME (hr, min)

Figure 2

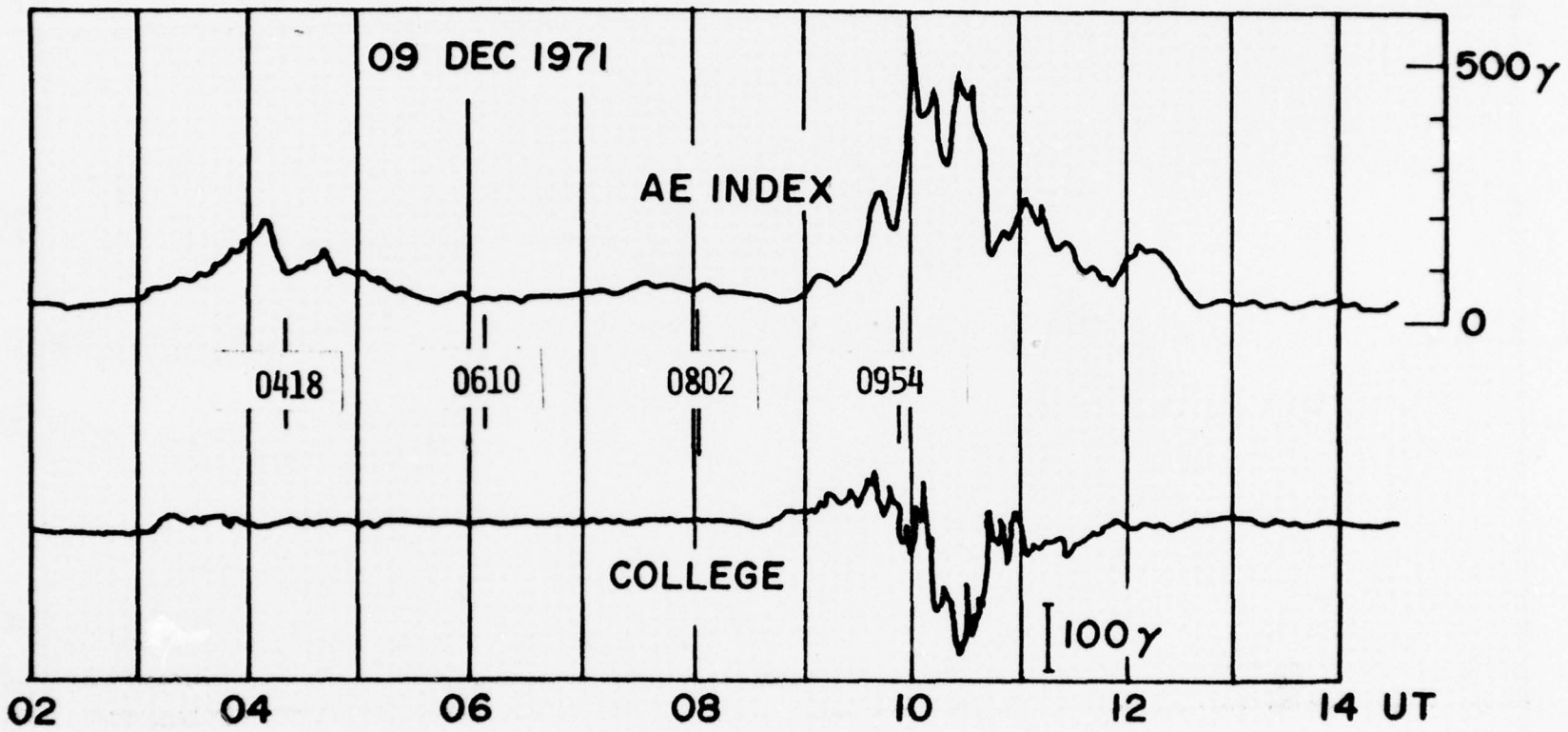

09 DEC 1971
AE INDEX
500γ
0
0418
0610
0802
0954
COLLEGE
100γ
02
04
06
08
10
12
14 UT

Figure 3

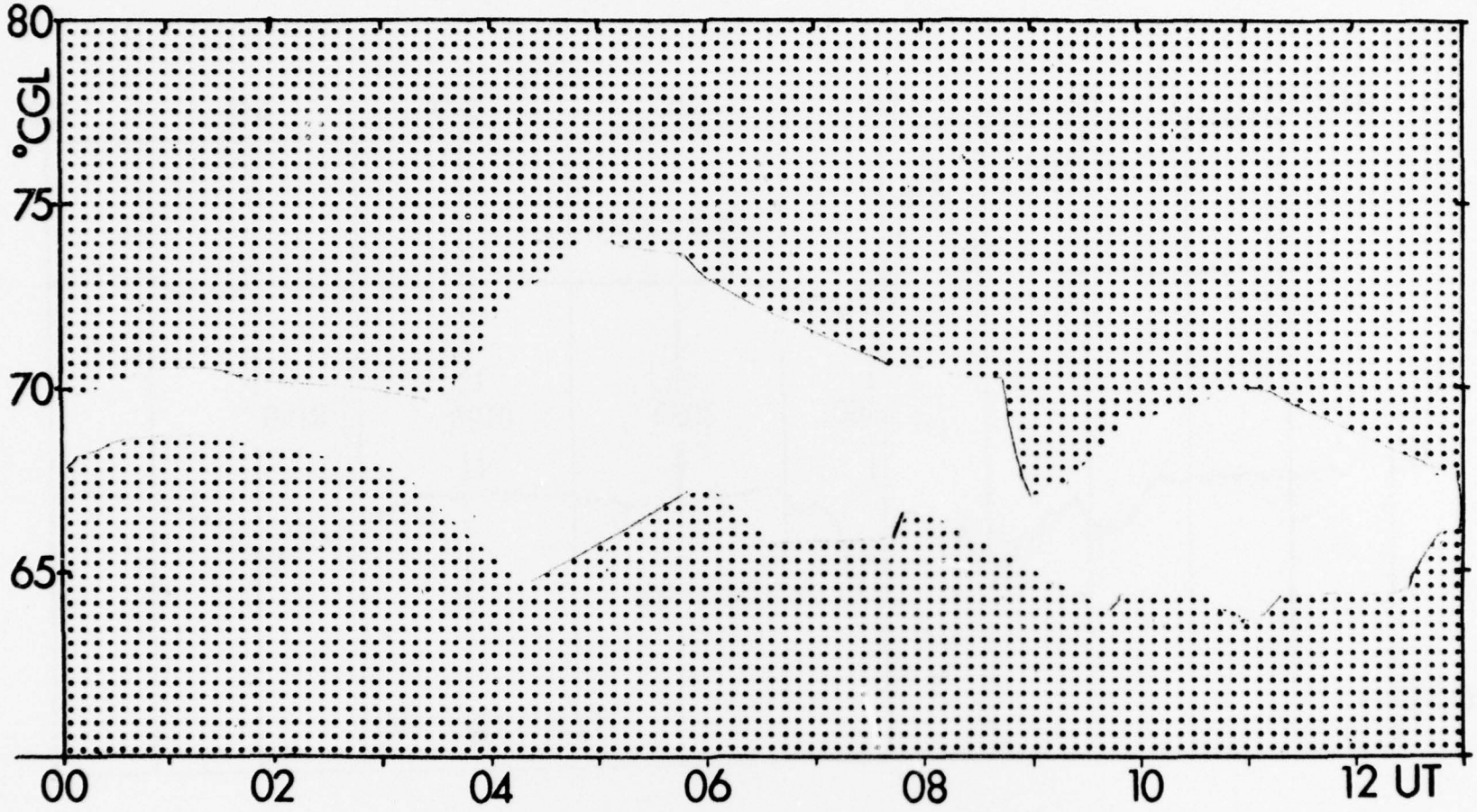
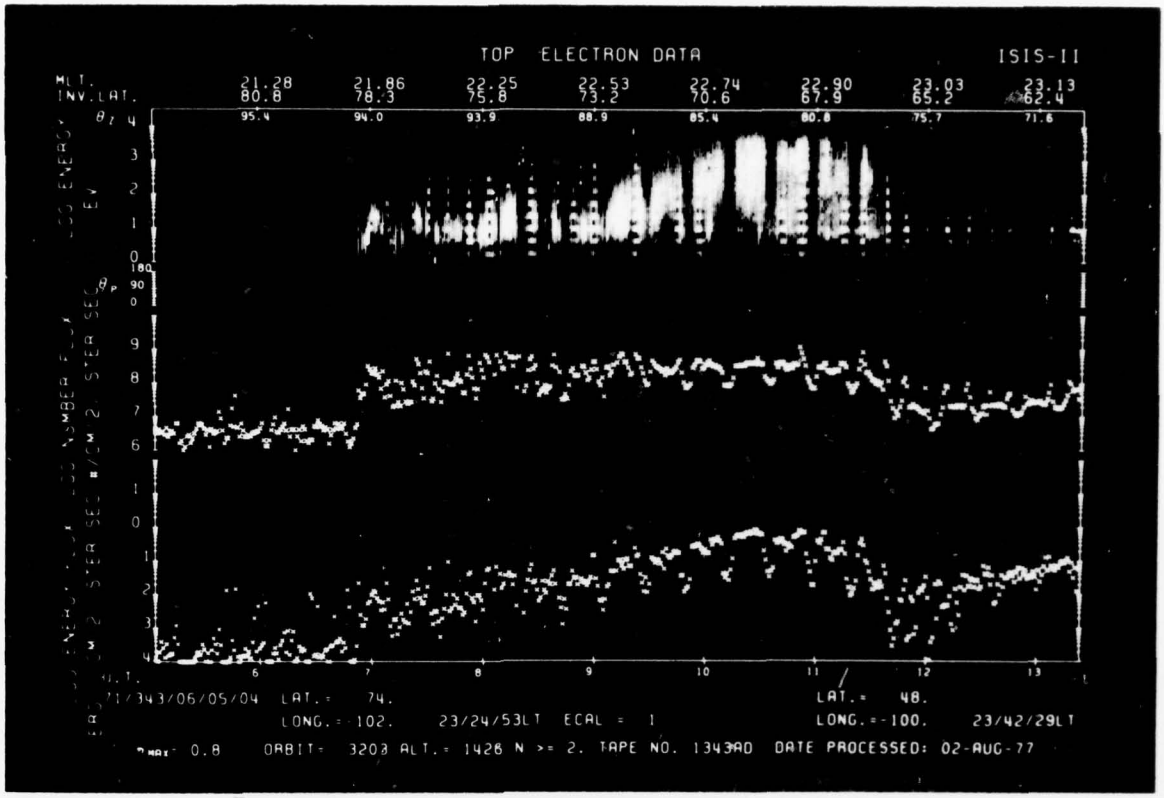


Figure 4


TOP ELECTRON DATA
ISIS-II
MLT.
INV.LAT.
21.28 80.8
21.86 78.3
22.25 75.8
22.53 73.2
22.74 70.6
22.90 67.9
23.03 65.2
23.13 62.4
95.4
94.0
93.9
88.9
85.4
80.8
75.7
71.6
LOG ENERGY EV
LOG ENERGY FLUX LOG NUMBER FLUX ERG/CM(2) STER SEC #/CM(2) STER SEC
U.T.
71/343/06/05/04 LAT.= 74.
LONG.= 102. 23/24/53LT ECAL = 1
LAT.= 48.
LONG.=-100. 23/42/29LT
ORBIT= 3203 ALT.= 1426 N >= 2. TAPE NO. 1343AD DATE PROCESSED: 02-AUG-77

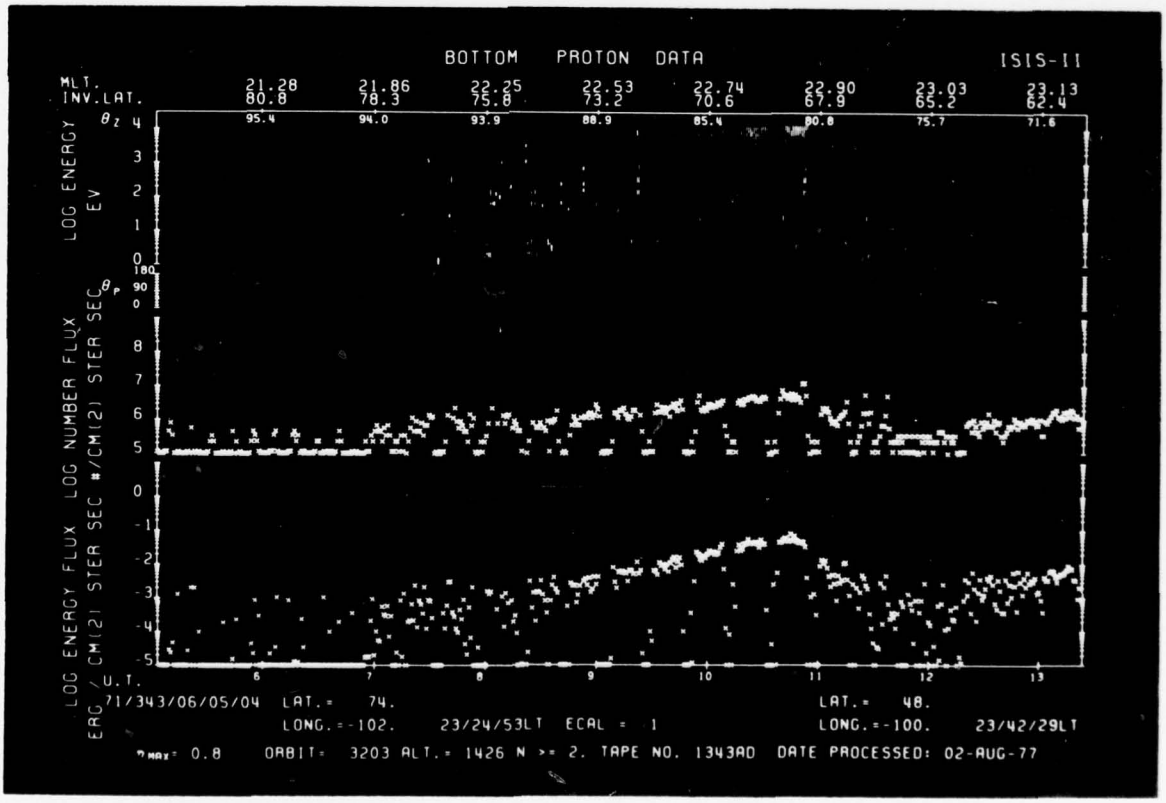
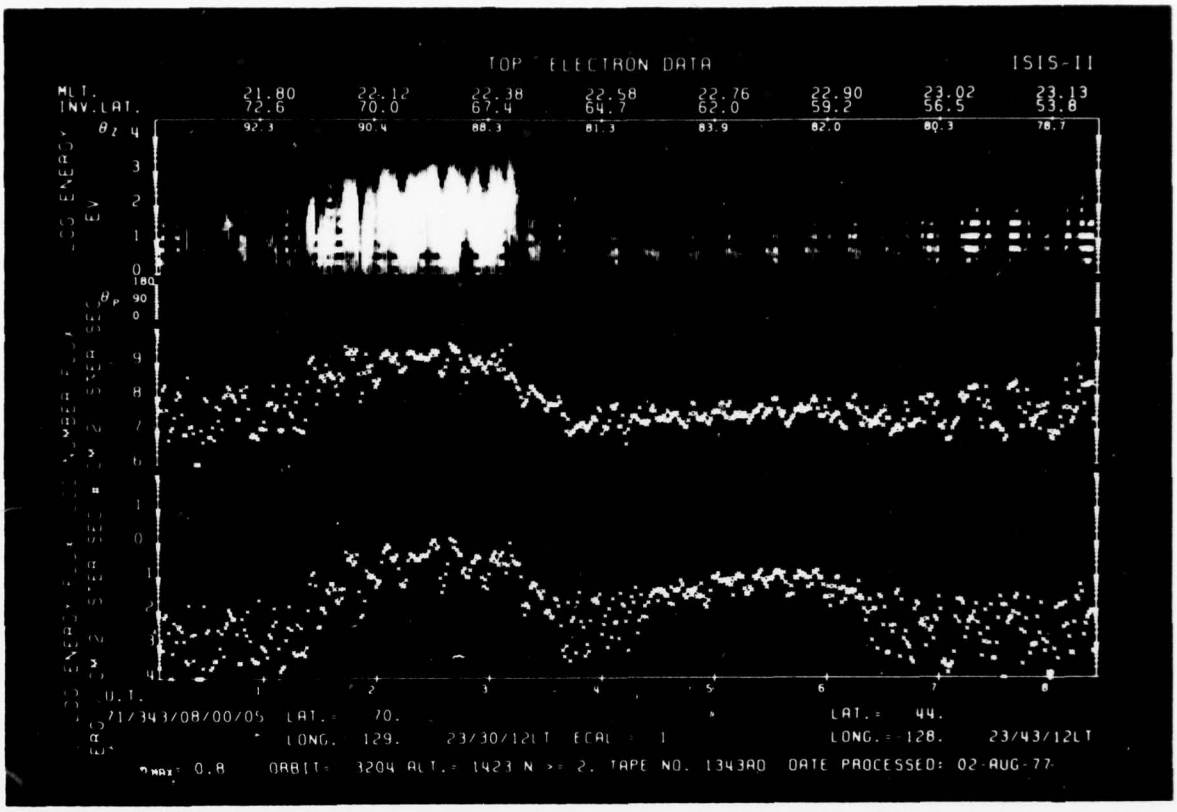

BOTTOM PROTON DATA
ISIS-II
MLT.
INV.LAT.
21.28 80.8
21.86 78.3
22.25 75.8
22.53 73.2
22.74 70.6
22.90 67.9
23.03 65.2
23.13 62.4
95.4
94.0
93.9
88.9
85.4
80.8
75.7
71.6
LOG ENERGY EV
LOG ENERGY FLUX LOG NUMBER FLUX ERG/CM(2) STER SEC #/CM(2) STER SEC
U.T.
71/343/06/05/04 LAT.= 74.
LONG.=-102. 23/24/53LT ECAL = 1
LAT.= 48.
LONG.=-100. 23/42/29LT
ORBIT= 3203 ALT.= 1426 N >= 2. TAPE NO. 1343AD DATE PROCESSED: 02-AUG-77

Figure 5


TOP ELECTRON DATA
ISIS-II
MLT.
INV.LAT.
21.80 72.6
22.12 70.0
22.38 67.4
22.58 64.7
22.76 62.0
22.90 59.2
23.02 56.5
23.13 53.8
92.3
90.4
88.3
81.3
83.9
82.0
80.3
78.7
LOG ENERGY EV
U.T.
71/343/08/00/05 LAT.= 70.
LONG. 129. 23/30/12LT ECAL = 1
LAT.= 44.
LONG.= 128. 23/43/12LT
ORBIT= 3204 ALT.= 1423 N >= 2. TAPE NO. 1343AD DATE PROCESSED: 02-AUG-77

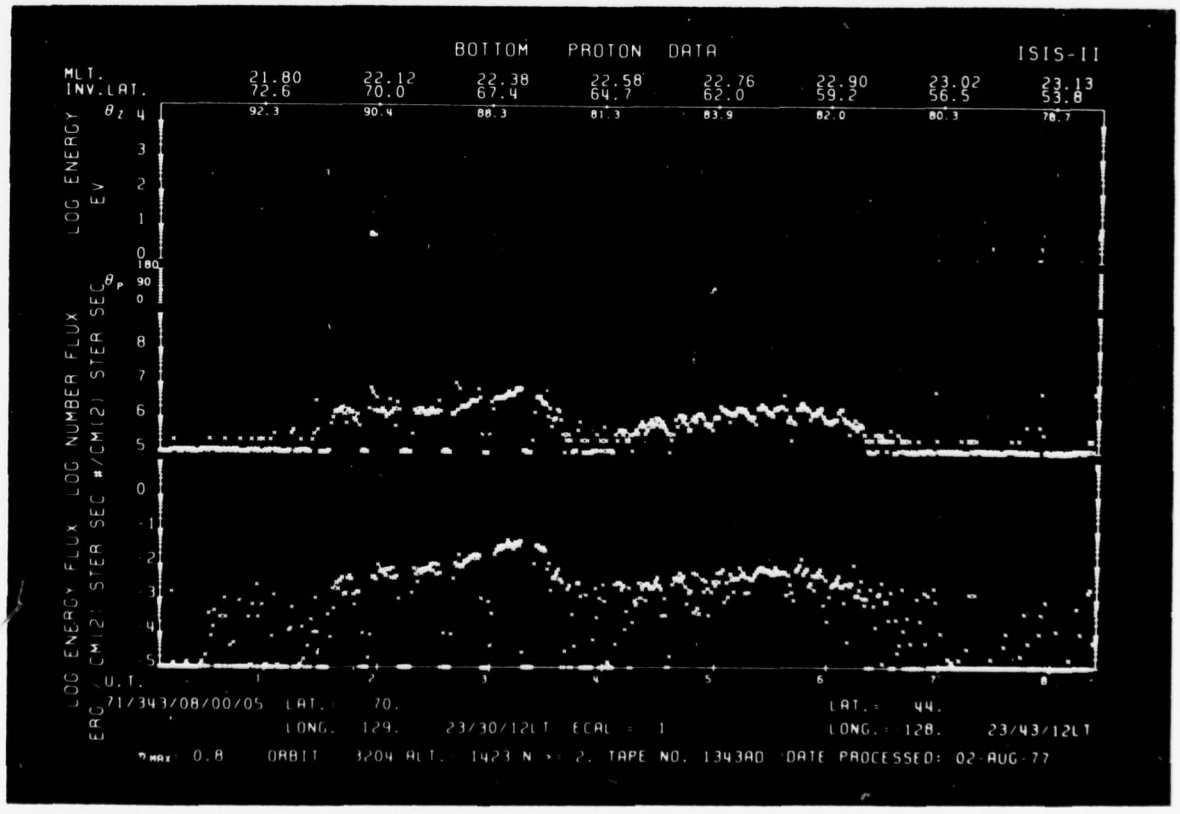

BOTTOM PROTON DATA
ISIS-II
MLT.
INV.LAT.
21.80 72.6
22.12 70.0
22.38 67.4
22.58 64.7
22.76 62.0
22.90 59.2
23.02 56.5
23.13 53.8
92.3
90.4
88.3
81.3
83.9
82.0
80.3
78.7
LOG ENERGY EV
LOG ENERGY FLUX LOG NUMBER FLUX ERG/CM(2) STER SEC #/CM(2) STER SEC
U.T.
71/343/08/00/05 LAT.= 70.
LONG. 129. 23/30/12LT ECAL = 1
LAT.= 44.
LONG.= 128. 23/43/12LT
ORBIT 3204 ALT. 1423 N >= 2. TAPE NO. 1343AD DATE PROCESSED: 02-AUG-77

Figure 6

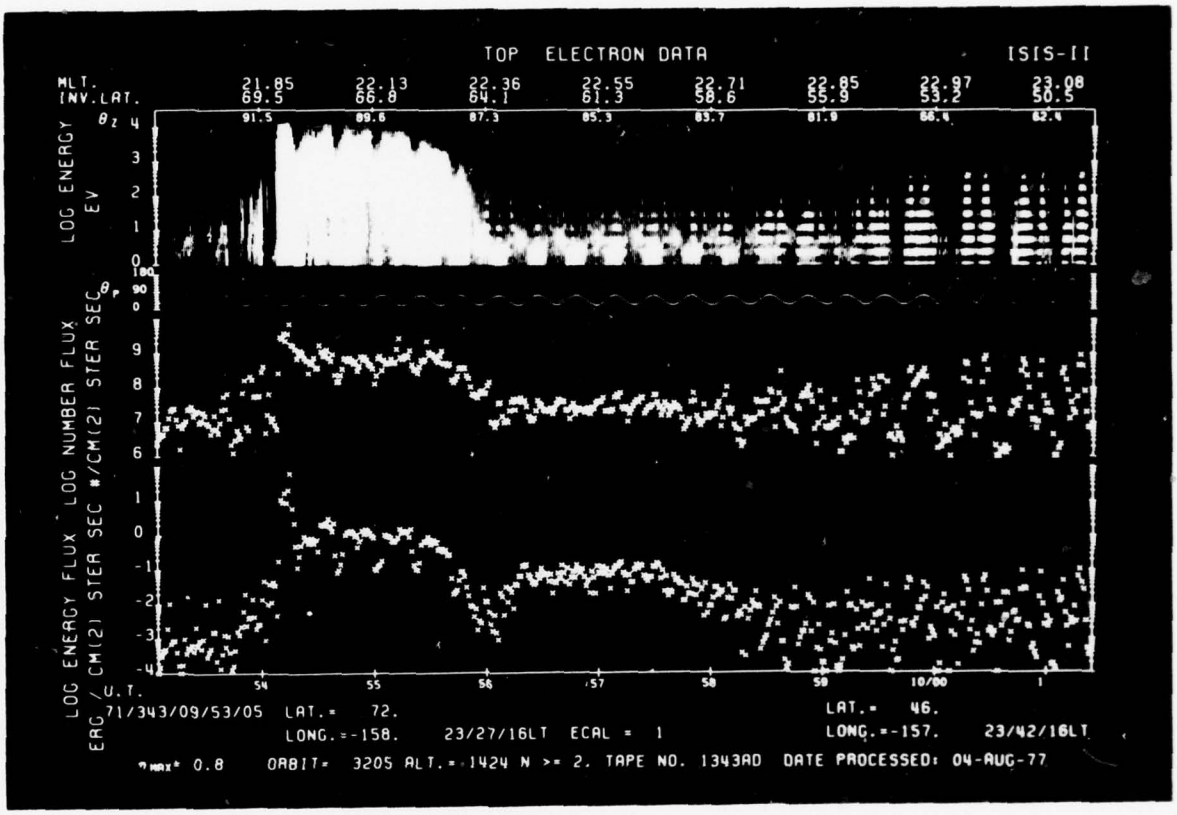

TOP ELECTRON DATA
ISIS-II
MLT.
INV.LAT.
21.85 69.5
22.13 66.8
22.36 64.1
22.55 61.3
22.71 58.6
22.85 55.9
22.97 53.2
23.08 50.5
91.5
89.6
87.3
85.3
83.7
81.9
86.4
82.4
LOG ENERGY EV
LOG ENERGY FLUX LOG NUMBER FLUX
EAG / CM(2) STER SEC #/CM(2) STER SEC
U.T.
54
55
56
57
58
59
10/00
1
71/343/09/53/05 LAT.= 72.
LONG.=-158. 23/27/16LT ECAL = 1
LAT.= 46.
LONG.=-157. 23/42/16LT
ORBIT= 3205 ALT.=1424 N >= 2. TAPE NO. 1343AD DATE PROCESSED: 04-AUG-77

Figure 7

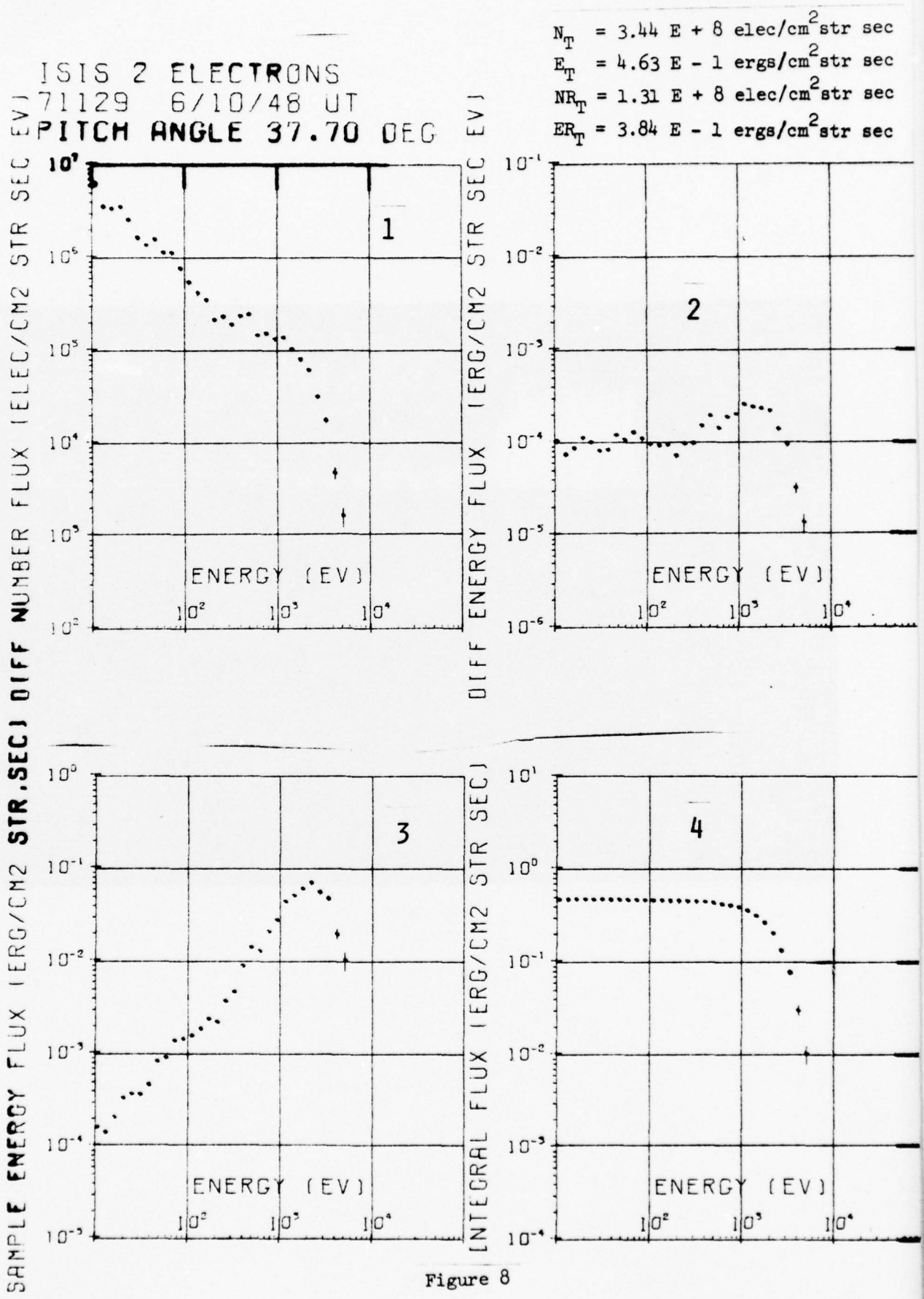
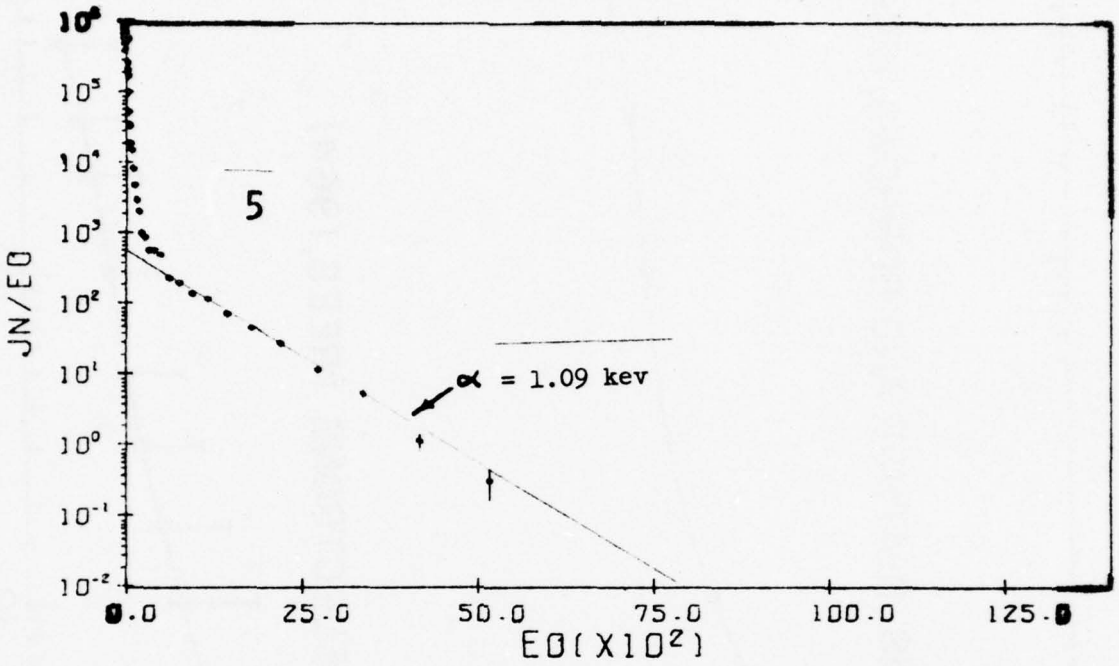

ISIS 2 ELECTRONS
71129 6/10/48 UT
PITCH ANGLE 37.70 DEG
N_T = 3.44 E + 8 elec/cm^{2}str sec
E_T = 4.63 E - 1 ergs/cm^{2}str sec
NR_T = 1.31 E + 8 elec/cm^{2}str sec
ER_T = 3.84 E - 1 ergs/cm^{2}str sec
DIFF NUMBER FLUX (ELEC/CM2 STR SEC EV)
DIFF ENERGY FLUX (ERG/CM2 STR SEC EV)
SAMPLE ENERGY FLUX (ERG/CM2 STR.SEC)
INTEGRAL FLUX (ERG/CM2 STR SEC)
ENERGY (EV)
1
2
3
4

Figure 8


ISIS 2 ELECTRONS
71129 6/10/48 UT
PITCH ANGLE 37.70 DEG
5
JN/E0
10^{6}
10^{5}
10^{4}
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
α = 1.09 kev
0.0
25.0
50.0
75.0
100.0
125.0
E0(X10²)

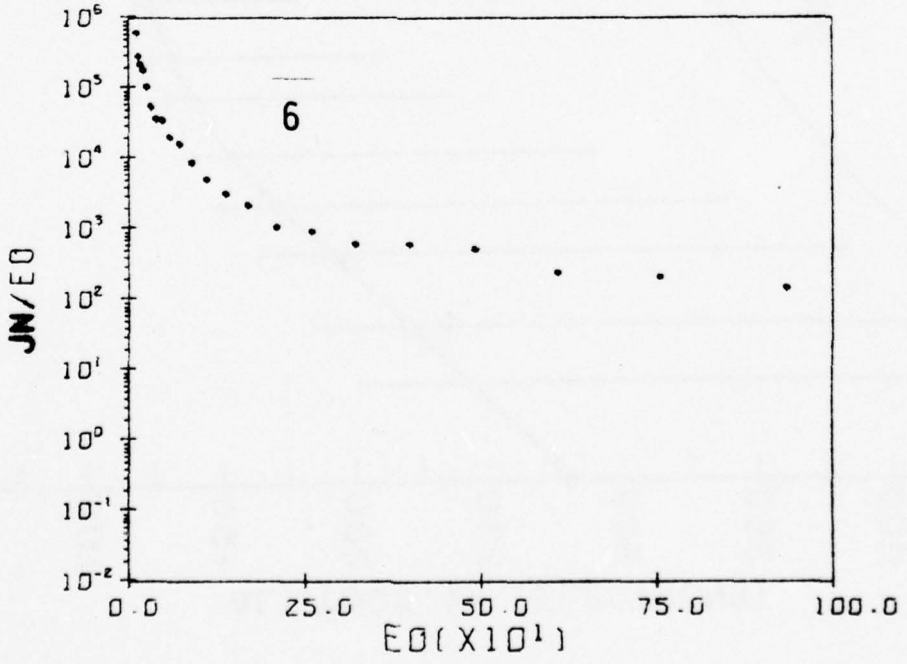

6
JN/E0
10^{6}
10^{5}
10^{4}
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
0.0
25.0
50.0
75.0
100.0
E0(X10¹)

Figure 8

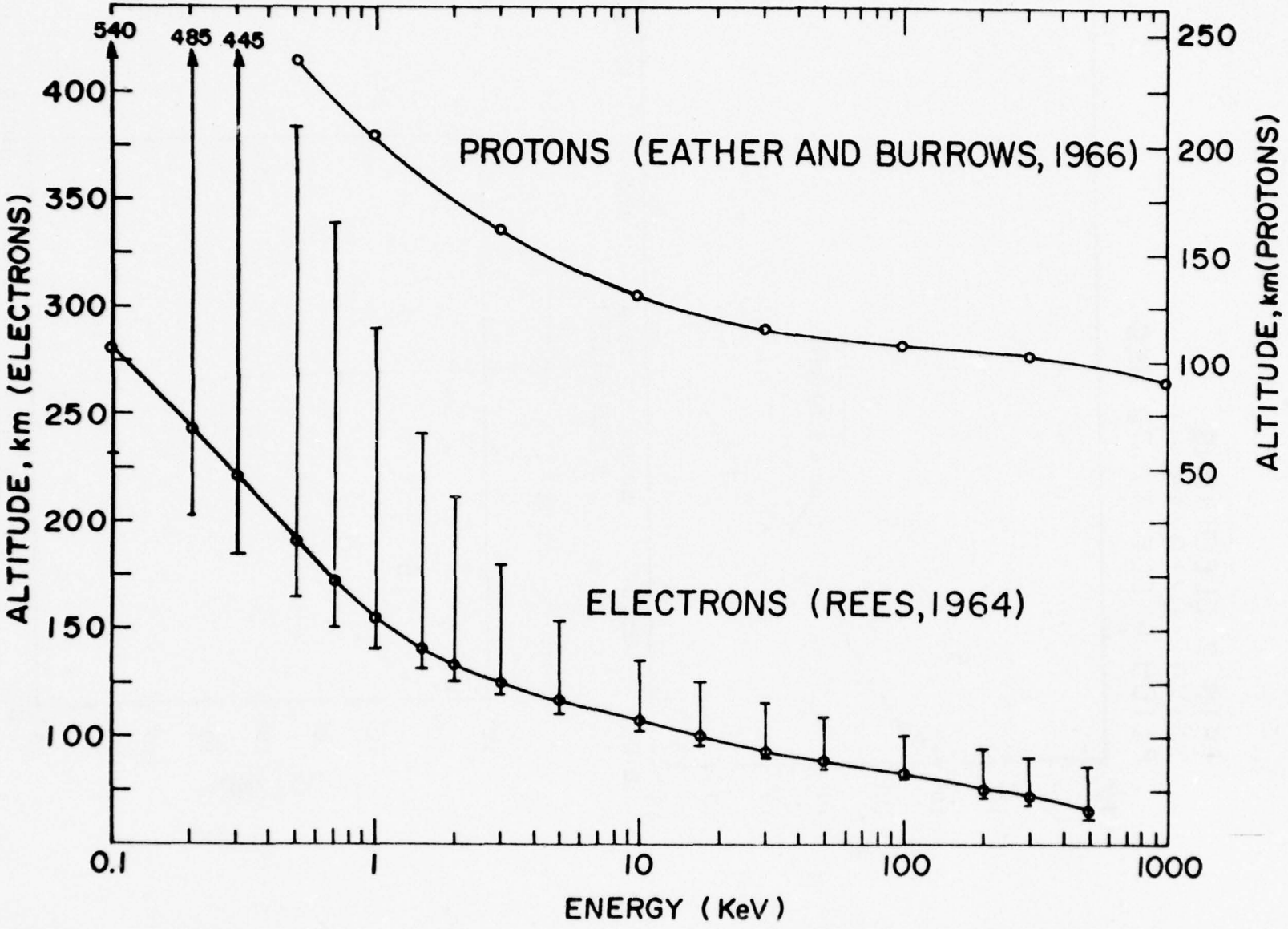

540
485
445
400
350
300
250
200
150
100
ALTITUDE, km (ELECTRONS)
PROTONS (EATHER AND BURROWS, 1966)
ELECTRONS (REES, 1964)
250
200
150
100
50
ALTITUDE, km(PROTONS)
0.1
1
10
100
1000
ENERGY (KeV)

Figure 9

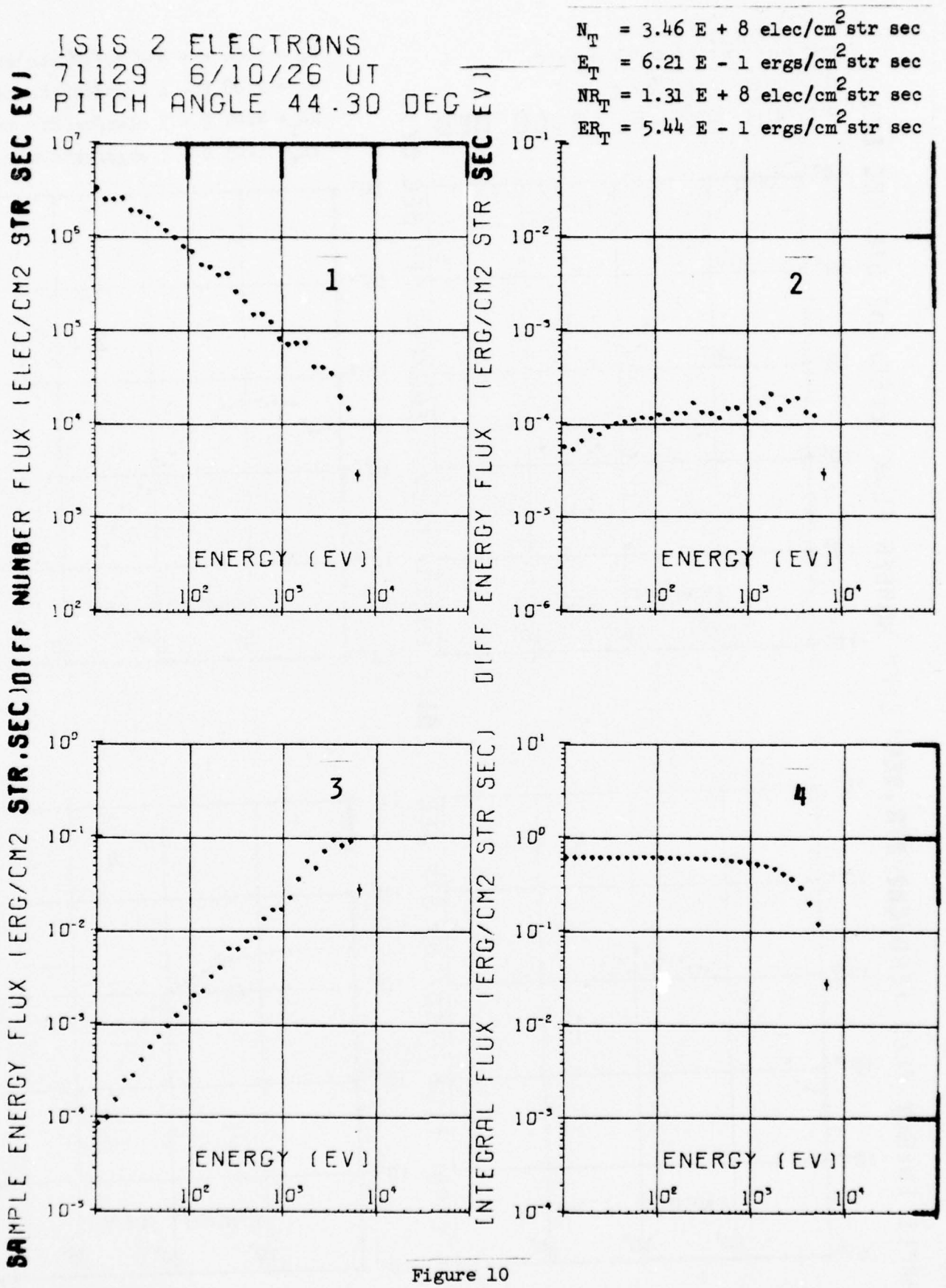
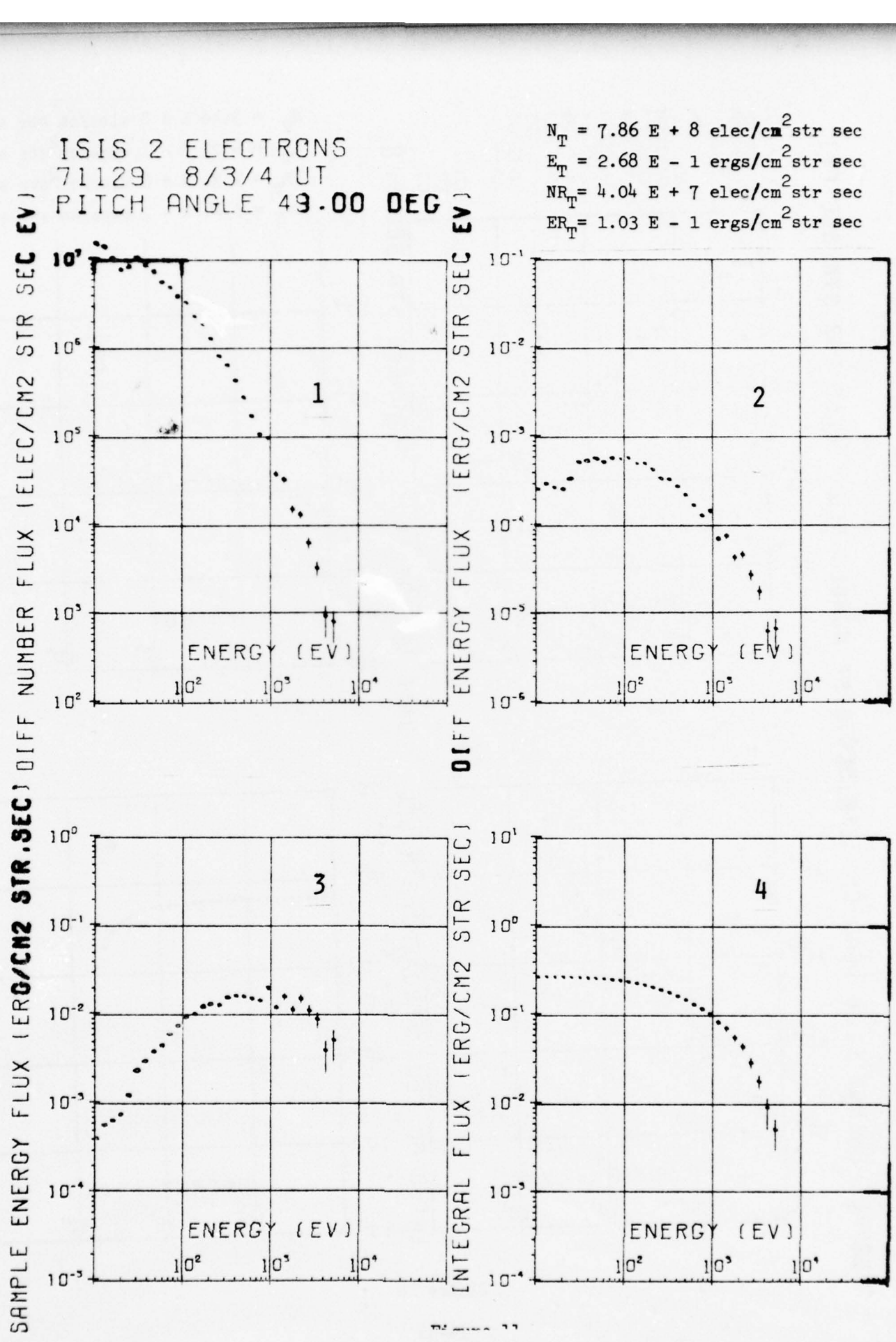

ISIS 2 ELECTRONS
71129 6/10/26 UT
PITCH ANGLE 44.30 DEG
N_T = 3.46 E + 8 elec/cm^{2}str sec
E_T = 6.21 E - 1 ergs/cm^{2}str sec
NR_T = 1.31 E + 8 elec/cm^{2}str sec
ER_T = 5.44 E - 1 ergs/cm^{2}str sec
DIFF NUMBER FLUX (ELEC/CM2 STR SEC EV)
DIFF ENERGY FLUX (ERG/CM2 STR SEC EV)
SAMPLE ENERGY FLUX (ERG/CM2 STR.SEC)
INTEGRAL FLUX (ERG/CM2 STR SEC)
ENERGY (EV)
1
2
3
4

Figure 10


ISIS 2 ELECTRONS
71129 8/3/4 UT
PITCH ANGLE 49.00 DEG
N_T = 7.86 E + 8 elec/cm^{2}str sec
E_T = 2.68 E - 1 ergs/cm^{2}str sec
NR_T = 4.04 E + 7 elec/cm^{2}str sec
ER_T = 1.03 E - 1 ergs/cm^{2}str sec
DIFF NUMBER FLUX (ELEC/CM2 STR SEC EV)
DIFF ENERGY FLUX (ERG/CM2 STR SEC)
SAMPLE ENERGY FLUX (ERG/CM2 STR.SEC)
INTEGRAL FLUX (ERG/CM2 STR SEC)
ENERGY (EV)
1
2
3
4

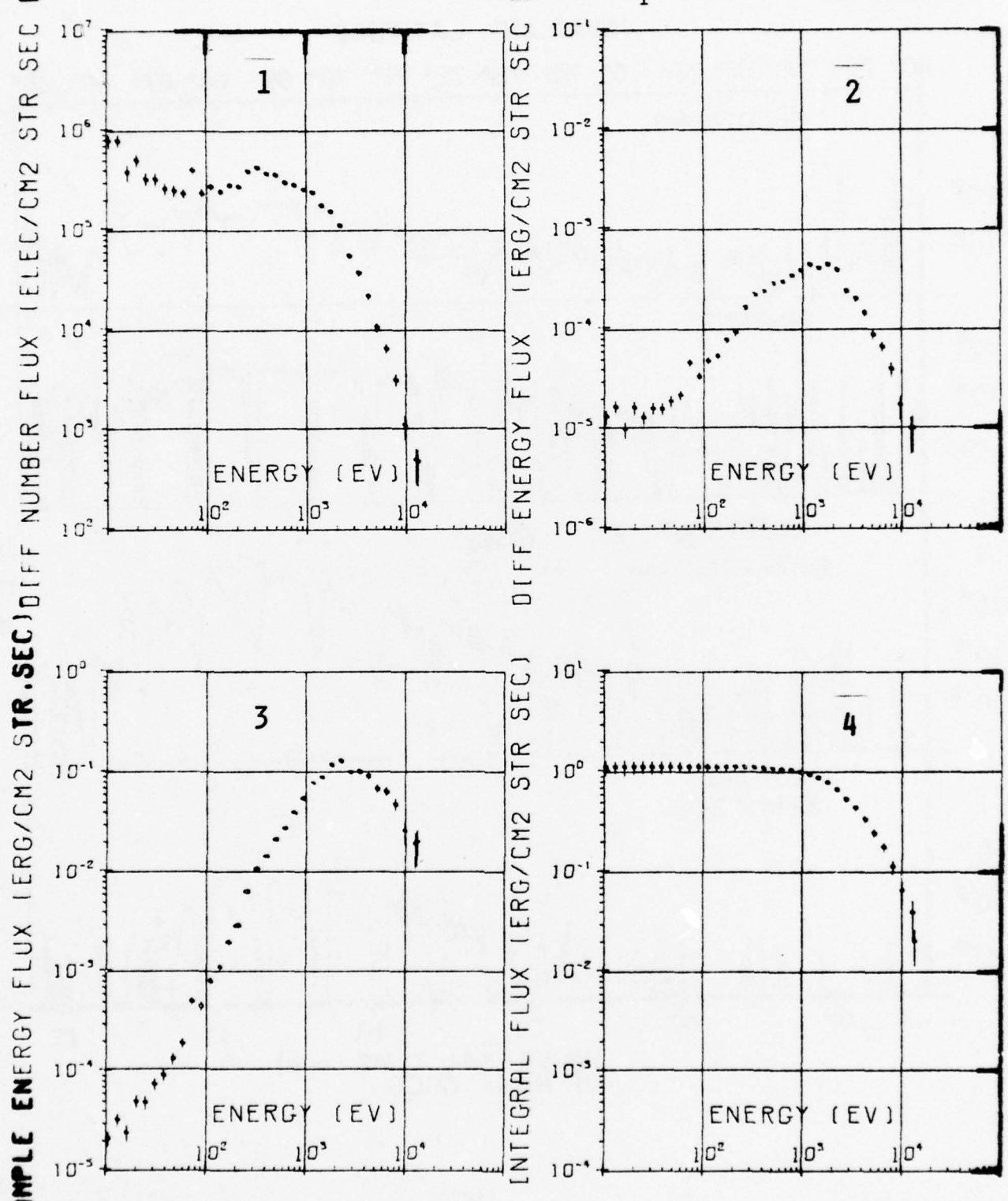

ISIS 2 ELECTRONS
71129 9/54/57 UT
PITCH ANGLE 44.00 DEG
N_T = 4.53 E + 8 elec/cm^{2}str sec
E_T = 1.11 E + 0 ergs/cm^{2}str sec
NR_T = 2.66 E + 8 elec/cm^{2}str sec
ER_T = 9.85 E - 1 ergs/cm^{2}str sec
1
2
3
4
DIFF NUMBER FLUX (ELEC/CM2 STR SEC EV)
DIFF ENERGY FLUX (ERG/CM2 STR SEC)
SAMPLE ENERGY FLUX (ERG/CM2 STR.SEC)
INTEGRAL FLUX (ERG/CM2 STR SEC)
ENERGY (EV)

Figure 12

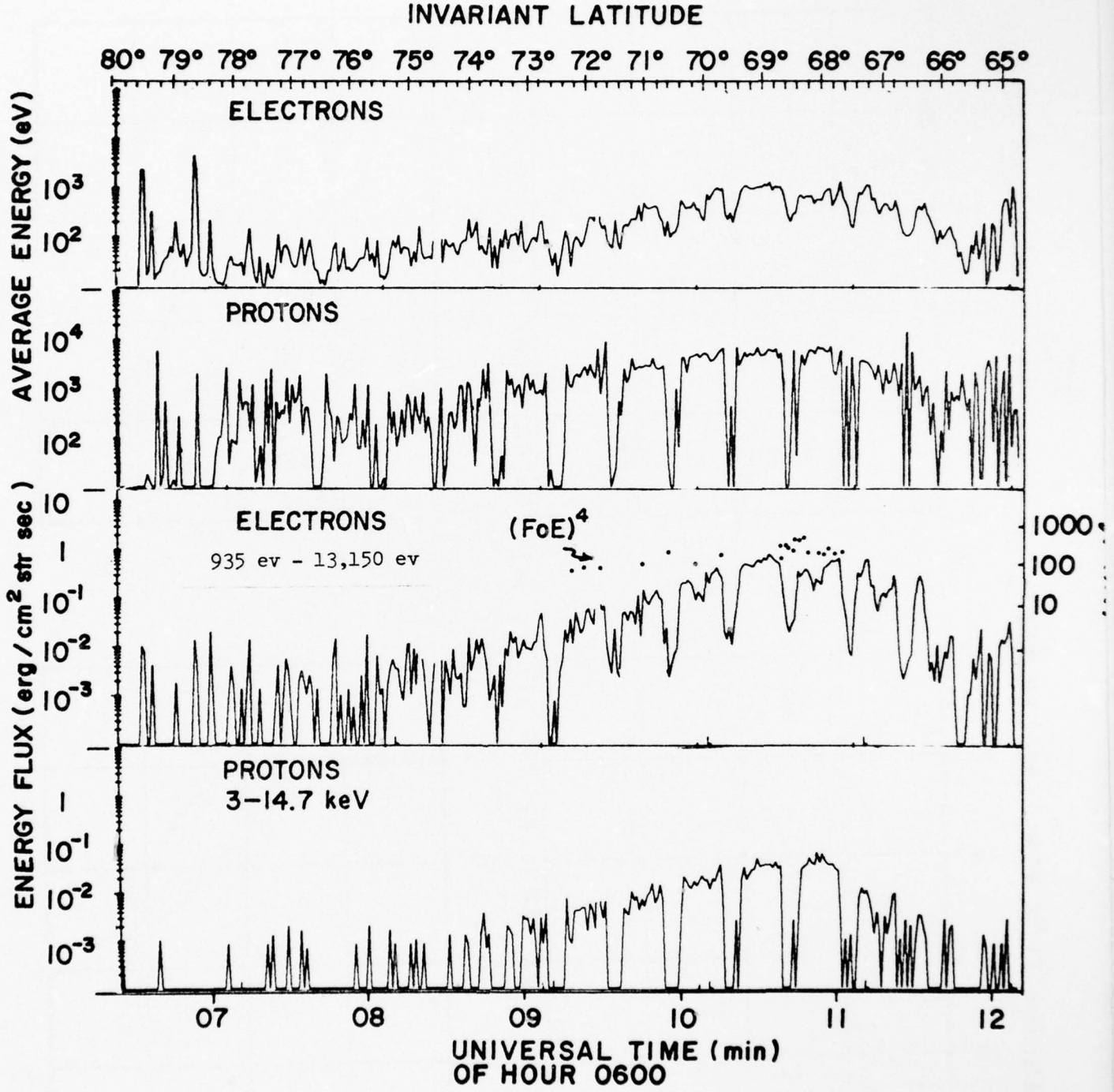

INVARIANT LATITUDE
80° 79° 78° 77° 76° 75° 74° 73° 72° 71° 70° 69° 68° 67° 66° 65°
ELECTRONS
PROTONS
ELECTRONS
935 ev - 13,150 ev
(FoE)4
PROTONS
3-14.7 keV
AVERAGE ENERGY (eV)
ENERGY FLUX (erg/cm2 str sec)
07 08 09 10 11 12
UNIVERSAL TIME (min)
OF HOUR 0600

Figure 13

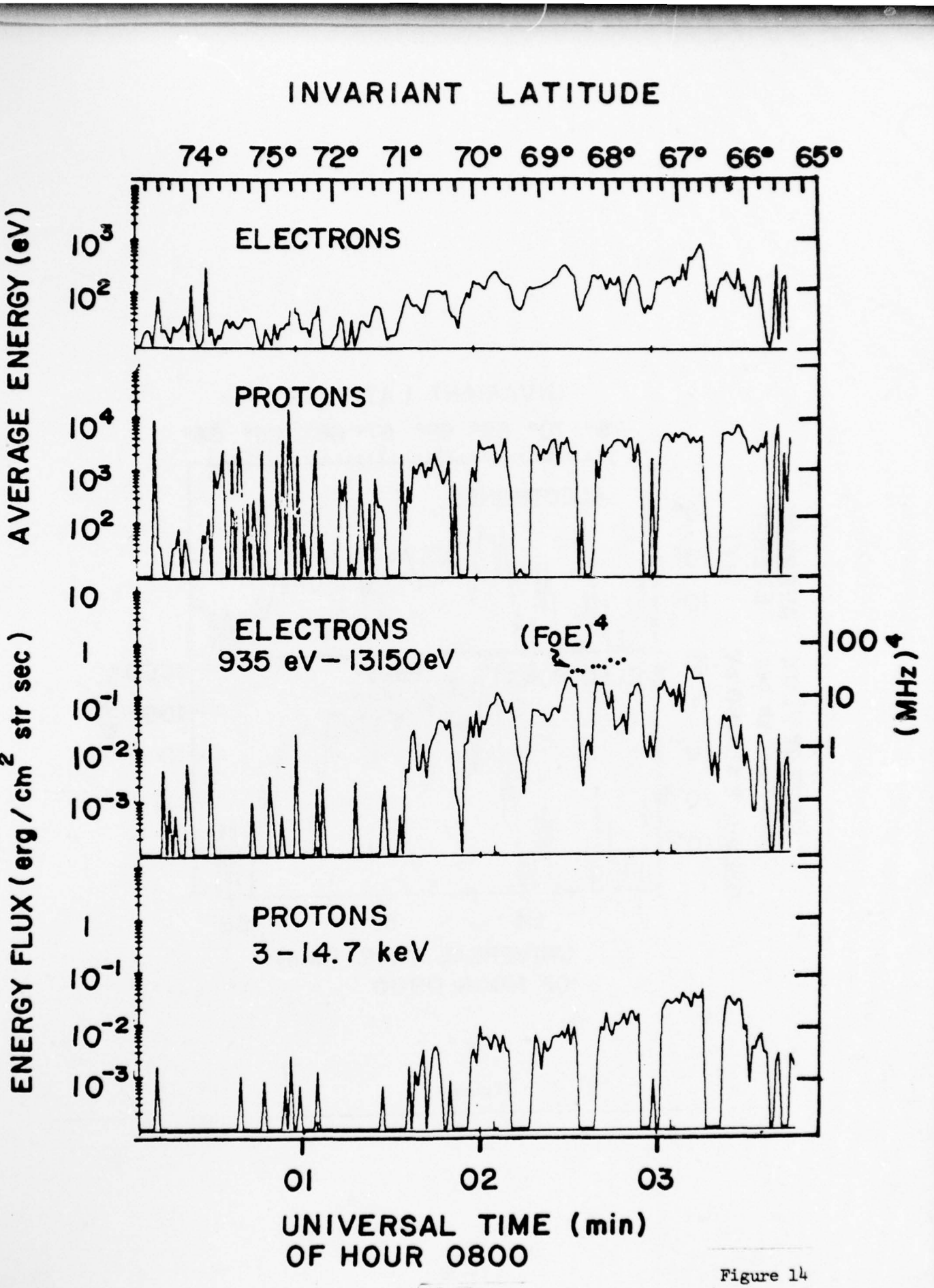

INVARIANT LATITUDE
74° 75° 72° 71° 70° 69° 68° 67° 66° 65°
AVERAGE ENERGY (eV)
ELECTRONS
PROTONS
ELECTRONS
935 eV – 13150eV
(FoE)4
PROTONS
3 – 14.7 keV
ENERGY FLUX (erg/cm2 str sec)
(MHz)
01
02
03
UNIVERSAL TIME (min)
OF HOUR 0800

Figure 14

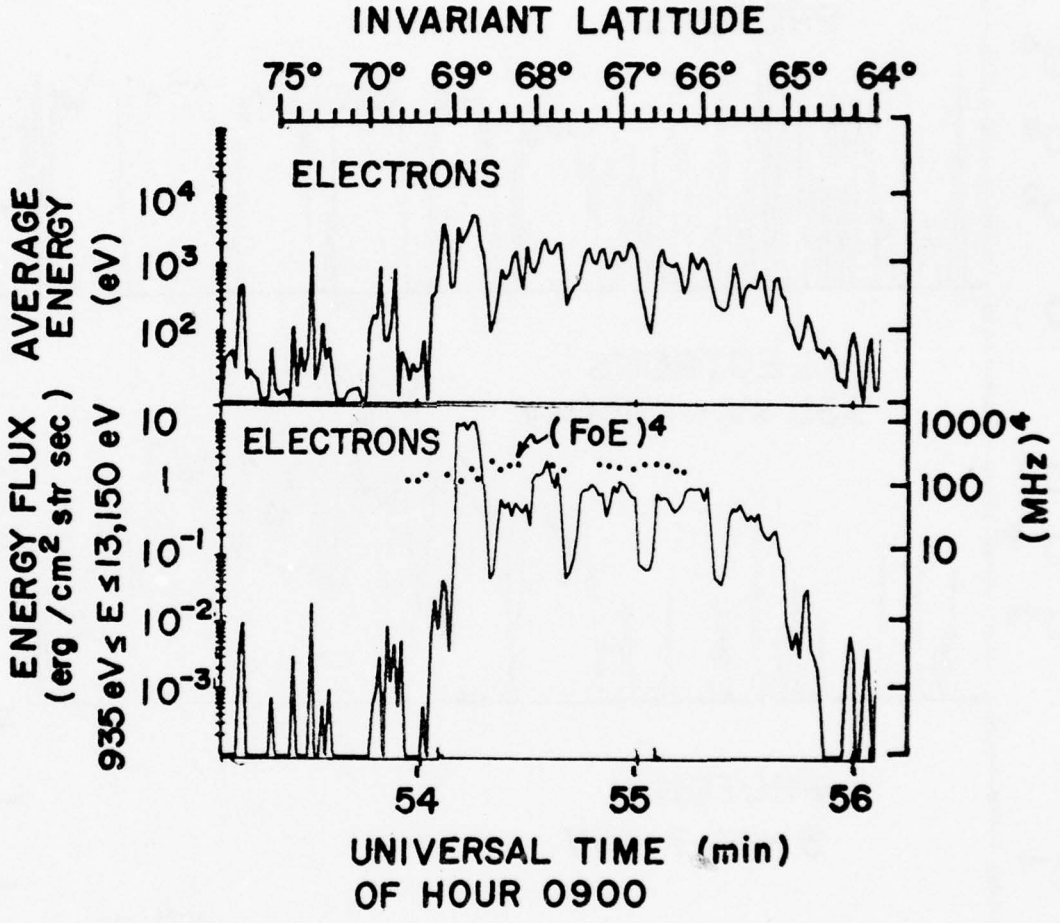

INVARIANT LATITUDE
75° 70° 69° 68° 67° 66° 65° 64°
ELECTRONS
AVERAGE ENERGY (eV)
10^4
10^3
10^2
ELECTRONS
(FoE)⁴
ENERGY FLUX (erg/cm² str sec) 935 eV ≤ E ≤ 13,150 eV
10
1
10^{-1}
10^{-2}
10^{-3}
1000⁴
100
10
(MHz)⁴
54
55
56
UNIVERSAL TIME (min)
OF HOUR 0900

Figure 15

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

DIRECTIONALLY SOLIDIFIED EUTECTIC MATERIALS

FOR ELECTRONIC, MAGNETIC AND OPTICAL APPLICATIONS

Prepared by:	Richard D. Sisson, Jr., Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Mechanical Engineering Worcester Polytechnic Institute

Assignment:	
(Air Force Base)	Hanscom AFB
(Laboratory)	Deputy for Electronic Technology
(Division)	Solid State Sciences
(Branch)	EM Technology
USAF Research Colleague:	Alton F. Armington, Ph.D.
Date:	August 19, 1977
Contract No.:	F44620-75-C-0031

DIRECTIONALLY SOLIDIFIED EUTECTIC MATERIALS FOR ELECTRONIC, MAGNETIC AND OPTICAL APPLICATIONS

by
Richard D. Sisson, Jr.

ABSTRACT

This report reviews the applications of directionally solidified eutectic material systems for electronic, magnetic and optical applications. The use of product effects in two phase material systems is briefly described. Product effects will allow the design of material systems for specific applications and properties.

Areas for possible future applications of eutectic systems are energy conversion, radiation detection and communications. Recommendations for future work are also made.

TABLE OF CONTENTS

INTRODUCTION

Directional solidification of a eutectic material produces an aligned two phase composite. These materials are sometimes called "in-situ" composites. The two phase structure yields properties that depend on the composition and microstructure of the material.

Composite materials properties may be related to a sum of the two phases individual properties, the anisotropy of the structure, the periodicity of the structure, the size of the phases in the structure or a product effect due to the coupling of the two phases (1). Structural applications of composites rely on the sum properties. The combination of a small amount of high strength phase with a low density matrix produces a material of increased strength and low density. Electrical, magnetic or optical applications of composites utilize the anisotropy, periodicity, size and product effects.

This report reviews the use of "in-situ" composites for electrical, magnetic or optical applications. New materials with unique properties may be developed for non-structural applications by the use of product effects.

OBJECTIVES

The objective of my summer research program was to review the use of "in-situ" composites for electrical, magnetic and optical applications. In addition, experimental work on cobalt-sulphur alloys for magnetic applications was initiated. An assessment of the possible application of eutectic or monotectic materials for microchannel plate fabrication has also been completed.

Eutectic Solidification

In a simple binary (or quasi-binary) system the composition with the lowest solidification temperature is the eutectic composition. This temperature is called the eutectic temperature. At the eutectic composition and temperature (eutectic point) the liquid phase is in equilibrium with two solid phases. If heat is removed from the system the liquid transforms into the two solid phases.

$$L \rightarrow \alpha + \beta$$

To illustrate a simple eutectic phase diagram the lead-tin system (2) is shown in figure 1. As shown in the figure the eutectic composition is 26.1 atomic percent lead and the eutectic temperature is 183°C.

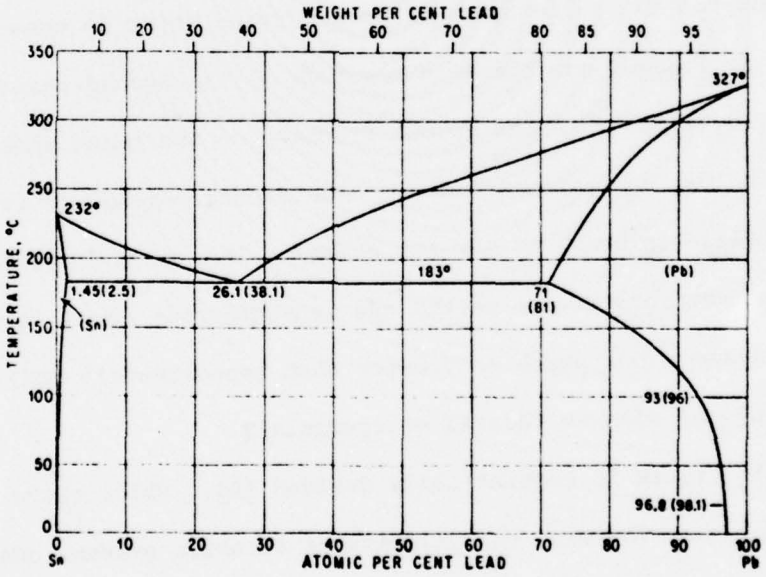


Figure 1. Lead-Tin Phase Diagram

Eutectic solidification produces a two phase alloy. Directional solidification can align these phases.

Directional Solidification

Directional solidification (also called unidirectional solidification) ideally consists of removing heat from a molten material in a single direction causing the planar solid-liquid interface to move through the material. Several directional solidification techniques have been developed to grow single crystals (3). These methods may also be used to produce aligned two phase structures in eutectic materials.

The procedure most easily applied to eutectic materials uses a chill block and a furnace. The chill block removes heat in one direction from the eutectic. Either the furnace or sample may be moved to advance the liquid-solid interface through the material. The result is an aligned two phase structure.

Eutectic Morphology

In general two basic types of morphologies exist in eutectics (4,5). The first is rods of one phase in a matrix of the second phase. This rod morphology develops when the volume fraction of the minor phase is small (i.e., less than approximately 28%). The second morphology is an alternating lamellae (plates) of the two phases. This morphology develops when the volume fractions of the phases are approximately equal (i.e., the fraction of the minor phase is greater than approximately 28%). These morphologies may also be faceted or irregular.

The 28% figure is theoretically derived (4). While there are contrary experimental observations, the majority of eutectic systems conform to the theoretical estimate.

The morphology also depends on the growth parameters. The important growth parameters are the growth velocity (R) and the temperature gradient in the solid phase at the solid-liquid interface (G_s)(2). In some systems

a lamellar structure breaks down to form rods at high growth rates (6,7).

The interlamellar or rod spacing (λ) is also a function of growth rate (R) with,

$$\lambda^2 R = \text{Constant}.$$

Directional Solidification of Eutectic Alloys

There are many general reviews of eutectic solidification (8,9,10, 11, to name a few). Kerr and Winegard (8) present the historical development and general theories. They also further develop the general theory of eutectic morphology (9). Hogan, Kraft and Lemkey (10) review the historical development as well as steady state growth, lamellar and fibrous (rod) morphologies, crystal orientation and eutectic classification. Thompson and Lemkey (11) reviewed the preparation techniques and properties of directionally solidified eutectic super alloys. May and Jones (12) present a short review and recommend the inclusion of eutectic solidification (preparation and properties) into the university teaching curricula. The review of the early work on the structure of eutectics was presented by Brady (13) in 1922.

There has been a great deal of experimental and theoretical work in eutectics. The major emphasis to date has been on structural applications. There have been many conferences as well as sessions at general meetings on eutectic solidification. A review of any abstract service will bring the reader up to date.

EUTECTIC MATERIALS FOR NONSTRUCTURAL APPLICATIONS

The application of eutectic materials for nonstructural applications was realized some years ago. Most of the early work was conducted at RCA, Siemens, United Aircraft, and Lehigh University. Galasso (14,15) reviewed the use of directionally solidified eutectics for optical, electronic and

magnetic applications in 1967. At that time he pointed out thermoelectric, magnetoresistive, infrared polarizing, superconducting, optical, ferromagnetic, and display applications; with the eutectic systems for each application. He stated that at that time over one hundred eutectic systems have been studied to some extent. Tables (15) were presented giving the eutectic composition, eutectic temperature, and microstructure for simple eutectic systems as well as complex oxides, complex halides, organics, and complex intermetallic eutectic systems. Salkind, Lemkey and George (16) also reviewed the nonstructural applications for aligned eutectics. May (17) has also reviewed the directional solidification of eutectics for "tough applications." In addition to the review of mechanical properties, thermal stability, and the development of materials for structural applications, he points out the possibility of coupling effects (product effects) for nonstructural applications. Salkind and Lemkey (18) also generally reviewed eutectic applications. Besides structural applications a few nonstructural uses are presented. In addition to the applications already mentioned the creation of porous materials by unidirectional solidification of a monotectic is described. A monotectic reaction differs from a eutectic reaction in the respect that instead of two solid phases forming simultaneously on cooling -- a solid and a liquid phase form simultaneously. The liquid phase will eventually solidify at some temperature below the monotectic temperature.

Weiss (19) critically reviewed the electromagnetic properties of eutectic composites. In this article he reviewed the work at RCA on the InSb-Sb system and the work at Siemens on the InSb-NiSb system. Magnetoresistive, thermomagnetic and optical properties are presented. This system has been well studied and has found the most applications of any system for nonstructural applications to date. A noncontacting current indicator for electric locomotive brake systems, a contactless push-button with field-plate and movable magnet

and an infrared detector have been manufactured using the InSb-NiSb eutectics.

Currently, Goodrum (20) is investigating directionally solidified eutectics for device applications. He postulates display, photovoltaic, superconducting, thermal fatigue monitoring, electron emitting, and magnetic device applications. He has also tabulated one hundred and fifty-two eutectic systems with their microstructure and properties. This work is being conducted under an Air Force contract with the Deputy for Electronic Technology.

In 1975 Weiss (21) reviewed the physical properties of in-situ composites. He discussed the properties in terms of sum and product effects (1). He pointed out that an artificial combination of two phases (mechanical combination) may be better than or as good as an aligned eutectic structure. Superconducting wires of 10μ NbTi in a Cu matrix as well as fiber optics glass technology are used as examples. He does state that the study of new in-situ composites may lead to new materials or new phases that may have important or interesting properties. The detection of X-ray by an in-situ composite is cited as a possible application.

Electronic Applications

The initial experimental investigation of electronic applications of directionally solidified eutectic materials was conducted at RCA by Liebmann and Miller (22) on InSb-Sb systems in 1963. Single crystal rods of Sb developed in a matrix of InSb. The electrical resistivity, the thermoelectric power, and the thermal conductivity were experimentally measured as a function of the structural arrangement of the phases. The values of resistivity and thermoelectric power were explained by electric circuit analogs. A decrease in thermal conductivity was observed that they explained by the phonon-phase boundary interaction. A large anisotropy

in properties was observed.

In the same year Weiss and Wilhelm directionally solidified the NiSb-InSb eutectic (23). They observed a large magnetoresistive effect in material with 1μ x 50μ NiSb rods in the InSb matrix. In addition, Paul, Weiss, and Wilhelm (24) reported an infrared polarizing effect with the same material. Later, Weiss (25) reported that directionally solidified NiSb-InSb could be used to measure magnetic fields (using the Hall effect). He also reported applications for controlling magnetic fields, polarization filters, and contactless variable resistors. In 1966, Weiss (26) reported additional applications for this material, including a modulator for converting d.c. currents and potentials into a.c. currents and as a magnetically controlled resistor (used in the brake system on the German Federal Railroad). In addition to the NiSb-InSb system, Siemens investigated many other III-V compound systems (see Table II in ref. 14).

The work at Siemens and RCA was followed by several other investigations into directionally solidified eutectic semiconducting materials. Elliot and Hiscocks (27) investigated the electrical and magnetoresistance properties of Cd_3As_2-NiAs eutectic. The properties did not compare favorably with the InSb-NiSb system. Umehara and Koda (28) investigated the structure and phase boundary energies in the InSb-MnSb, InSb-NiSb, InSb-FeSb and InSb-CrSb systems.

Investigations were also conducted on silicon- and germanium-based eutectics (29,30,31). Hilbren and Hiscocks (29) found that the structures of $NbSi_2$-Si, $TaSi_2$-Si and $TiGe_2$-Ge produced structures that may have electronic applications. Levinson (30) determined the electrical and thermal properties of directionally solidified $CrSi_2$-Si eutectic. He observed a large anisotropy in the physical properties.

Hellawell (31) has reviewed silicon- and germanium-based eutectics. Although electrical properties are not presented, the structures of a large number of eutectics are reviewed.

Lamellar eutectics may also be used to create a multilayer P-N structure. Albers and Verberkt (32) have pointed out the possibility of developing alternating layers of P-type and N-type semiconductor via directional solidification of eutectic materials. They have directionally solidified the SnSe-$SnSe_2$ eutectic to produce a P-N multilayer structure. Currently, Yue at UCLA (33) is investigating the SnSe-$SnSe_2$ system to produce a P-N multilayer structure for photovoltaic energy conversion.

Favier et al (34) investigated the superconducting properties of selected eutectics. The Pb-Sn, $AuSn_4$-Sn, and In_2Bi-In systems were investigated. Magnetization measurements were made. The Josephson effect was also studied.

Magnetic Applications

There has been a substantial amount of work on the use of directionally solidified eutectic materials for magnets. This work is a logical extension of the magnetic materials fabrication technology used to produce the Alnico series of permanent magnet alloy steels (35,36). Alnico steels contain Fe, Al, Ni, Cu, Co and Ti in various proportions. The magnetic properties of these steels are enhanced by preferred grain orientation and by cooling in a magnetic field. A spinoidal decomposition phase transformation in the magnetic field produces an aligned two phase structure. This aligned mixture creates a magnetically anisotropic material. The reason for the enhanced magnetic properties is that the size of the second phase particles (approximately 300 x 300 x 1200Å) creates single-domain particles. This increases the magnetic coercivity of this material. There

is a shape as well as a crystallographic anisotropy contribution to magnetic properties.

The single-domain particle principle has been applied to several ferromagnetic eutectic systems. Albright and Kraft (37) investigated the Fe-FeS eutectic. They observed rods of Fe in the FeS matrix. An increase in the coercive force from less than one oersted for iron to ten to twelve oersteds for the directionally solidified eutectic alloy was observed (38).

Galasso et al (39) investigated the magnetic properties of Fe-FeSb eutectics. This system produced rods of Fe in the FeSb matrix. However, a larger volume fraction of Fe rods was expected based on the phase diagram. They determined a coercive force of between twelve and eighteen oersteds for the directionally solidified materials that was dependent on Fe rod size (between 1 and 2μ). The fine rod size (approximately 1μ) produced the largest coercive force.

Galasso et al (40) have directionally solidified the $BaFe_{12}O_{19}$-$BaFe_2O_4$ eutectic. This system was selected because of the magnetic properties of $BaFe_{12}O_{19}$. Blades (lamellae) of the magnetic $BaFe_{12}O_{19}$ were observed in the $BaFe_2O_4$ matrix. No magnetic measurements were reported because the samples obtained were too small. Problems were also encountered because the C axes (axis of easy magnetization) of the $BaFe_{12}O_{19}$ were not aligned. Therefore, no attempts were made to produce larger samples for magnetic measurements.

Jackson, Tauber and Kraft (41) investigated the magnetic anisotropy of the MnSb-Sb aligned eutectic. Two-micron rods of ferromagnetic MnSb were observed in the Sb matrix. The anisotropic magnetic behavior was found to be due to opposition of the effects of shape anisotropy and crystallographic anisotropy. In addition, they reported that strain factors

may also be important. Yim and Stofko (42) observed a large magnetic anisotropy in Bi-MnBi single crystal eutectics. Needles of ferromagnetic MnBi were observed in the Bi matrix. The axis of easy magnetization (C axis) was the long axis of the needles. The magnetic moment, the Hall coefficient, the electrical resistance, the magnetoresistance, the Seebeck coefficient, and the thermal conductivity were determined.

Colling and Kossowsky (43) investigated the soft magnetic properties of Co-Nb and Co-Nb-Fe directionally solidified composites. Fe alloy additions lowered the coercive force, increased the magnetization, and lowered the Curie temperature. Lamellae of $NbCo_3$ were in a Co rich matrix.

Livington (44) determined the structure and magnetic properties of the Au-Co aligned eutectic. Co rods were observed in an Au matrix. The coercive force was found to increase with growth rate (finer Co rods) to 330 oersteds at .21 cm/sec. A subsequent wire drawing increased the coercive force to 920 oersteds (due to the even finer rod size). Sahm and Killias (45) also directionally solidified Au-Co eutectic composites. The directional solidification was performed in a magnetic field. Structural changes were reported due to the magnetic field. No magnetic properties were reported.

In addition, Galasso (14) has reported on other magnetic eutectics including Co-Y_2Co_{17}, Co-CoSb, and Ni-Ni_3Sn. Structural and magnetic data are reported. Enhanced anisotropic magnetic properties were observed in these systems.

Optical Applications

The optical properties of eutectics are a function of the size and periodicity of microstructure. In addition, the absorption edge of the phases on the microstructure may also be utilized.

The systems most widely studied are alkali halide mixtures. Lotham and Hellawell (46) investigated the fluorides, chlorides, and bromides of lithium, sodium and potassium. Of the thirty-six pairs investigated, twenty-one exhibited at least one eutectic point, six had two eutectics. Eutectic morphologies were determined on directionally solidified mixtures. Penfold and Hellawell (46a) further investigated the microstructures of the LiF-NaF and NaF-NaCl eutectic systems.

Batt, Douglas and Galasso (47) determined the optical properties of the directionally solidified NaCl-NaF eutectic. Image transmission properties similar to fiber optics materials could be obtained. They also reported that this directionally solidified eutectic was a far-field infrared transmitting medium for wavelengths larger than the inter-NaF rod distance (approximately 6μ).

Yue and Yu (48) also explored the optical properties of the NaCl-NaF eutectic. A model was presented to explain the phenomena of transmittance versus far-field infrared wavelength in the eutectic. The data are in excellent agreement with the theory.

Sievers (49) reported on the infrared Bragg scattering of NaCl-NaF eutectic. The results were explained in terms of the eutectic microstructure and optical theory.

Bright and Lewis (50,51) investigated the LiF-MgF_2 system. They reported the microstructure and crystallographic relation between the phases. They also reported on the NaF-MgF_2 system. The $NaMgF_3$ compound was observed to form a eutectic with NaF and MgF_2.

van Hoof and Albers (52) have reported the conversion of X-rays into visible light in the NaCl-$PbCl_2$ eutectic. X-ray excited luminescence of NaCl had a maximum at 30-40 volume percent $PbCl_2$. The results were explained

in terms of the eutectic product properties.

Davies, Clauson and Wieder (53) investigated the infrared polarizing properties of InSb-Sb eutectic films. The Sb lamellae in the InSb matrix acted as a grating to polarize the infrared transmitted beyond the InSb absorption edge. Paul et al (24) have also reported on infrared polarization in InSb based eutectics. The optical properties of the CeSb-InSb, MnSb-InSb and FeSb-InSb were determined.

Superconductivity in Eutectics

In 1963, Cline, Rose and Wulff (54) investigated the superconducting properties of columbium (niobium) -thorium eutectic. They observed a fortyfold increase in the critical current carrying ability of the eutectic as compared to columbium wire. Their alloys were not directionally solidified. Levy, Kim and Kraft (55) investigated the superconducting behavior of Cd-Sn directionally solidified eutectics. They observed no improvement in properties over bulk phases.

Refractory Oxide Eutectics

The largest study of refractory oxide-metal eutectics was conducted by Chapman et al (56,57,58) on the UO_2-W system. They observed rods of conducting W in a UO_2 matrix. This material may have structural applications with the high modulus fibers in the oxide matrix. There may also be electrical and thermal applications.

One possible application of this system proposed by Goodrum (59) is a cold cathode. This uses the electron emission for the W fibers in the matrix to produce an electron beam.

Hulse and Batt (60,61) have explored the use of a floating molten zone technique to produce oxide eutectics. They present results for twenty-eight oxide eutectics.

Schmid and Viechnicki (62) reported on the Al_2O_3-ZrO_2 eutectic for high temperature applications. ZrO_2 rods were observed in an Al_2O_3 matrix.

SOLID-STATE DIRECTIONAL TRANSFORMATIONS

Eutectoid decomposition is a solid state phase transformation in which one solid phase transforms to two different solid phases on cooling.

$$\alpha \rightarrow \beta + \gamma.$$

Aligned phase structures can be obtained by directionally transforming a eutectoid material. The structure of the two phase mixture will be dependent on the structure of the high temperature single phase as well as the variables of transformation rate and temperature gradients.

Livingston (63) reviewed directional transformations in eutectoid alloys. He summarizes the results of ten directionally transformed eutectoid alloys. Carpay (64,65) has also reviewed aligned composites produced by directional eutectoid transformations. The eutectoid systems investigated include Cu-Cu_9Al_4, Cu-Cu_9In_4, NiIn-Ni_2In_3, Co-Co_2Si, Fe-Fe_3C, FeAl-$FeAl_2$ and Al-Zn.

In addition, Carpay (66) reported that in-situ composites can be obtained from decomposition of the vitreous state. Results in Li_2O-B_2O_3 are described. The $Ca(NO_3)_2$-$RbNO_3$, $ZnCl_2$-KCl, Bi_2O_3-B_2O_3, Bi_2O_3-GeO_2 and Li_2O-B_2O_3 have been successfully controlled. This technique offers greater flexibility in alloying than the eutectic or eutectoid transformations alloy. However, you must use glass-forming compounds.

RECOMMENDATIONS

1. Materials with unique and interesting properties may be obtained by product effects in two-phase mixtures. In his interesting paper, van Suchtelen (1) presents the concept of product properties. These product

properties are a result of the coupling of the two phase. For example, a piezomagnetic material coupled with a magnetoresistive material produces a product property of piezoresistance.

van Suchtelen presents a six by six matrix of properties (mechanical, magnetic, electrical, optical, thermal and chemical). He explains that any effect can be represented by an input of one the variables, producing an output of the second. (e.i., mechanical input → electrical output is piezoelectricity or piezoresistivity; optical input → magnetic output is photomagnetic effects.) The coupling then of two phases can produce unique properties.

The materials that may produce the most interesting effects when coupled and aligned are the ferroelectric type compounds. These materials often display piezo-effects and pyro-effects, etc. The materials of interest are the titanates and ferrites.

van der Boomgaard et al (67,68,69) directionally solidified $BaTiO_2$-$CoFe_2O_4$ eutectics. This system was chosen based on van Suchtelen's predictions. This perovskite-spinel structure is used to convert magnetic and electric signals. The combination of electrostriction in phase 1 with piezomagnetism in phase 2 produces this effect. The coupling is mechanical.

In some cases these effects can be obtained without directional solidification. A non-aligned two phase mixture may produce the same result. In some case the alignment of the phases will enhance the effect.

I recommend that eutectic systems be explored based on predictions from van Suchtelen's matrix.

2. Several eutectic systems have been explored for superconducting properties (34,54,55). The Nb_3Ge composition has been reported to be superconducting at temperatures up to 20.4°K (70). This material was produced

by chemical vapor deposition. Investigation into eutectics with this system may be of interest.

3. Currently work is in progress at Georgia Tech(59) on the use of UO_2-W eutectics as a cold-cathode for electron emission. Lanthanum hexaboride is a more efficient electron producer than tungsten. Eutectics with lanthanum hexaboride should be explored for cold-cathode applications (71).

4. Directionally solidified eutectics or monotectics may be useful for radiation detection or sensing (72). Using the principles applied to electron multiplication in channel electron multipliers, radiation (i.e. β^-, γ, infrared) may be detected and amplified. See ref. 72 for details.

5. Solar energy conversion is a topic of great interest today. Directionally solidified eutectics may be used as P-N materials for photovoltaic energy conversion. Currently Yue (73) at UCLA is investigating the SnSe-$SnSe_2$ system for this application. Other materials systems have higher conversion efficiencies than this system. I recommend work be initiated on eutectic systems with GaP, CuS, and CdS for higher energy conversion efficiencies.

Another method of converting solar energy into a useful form in the photoelectrolysis of water to produce hydrogen and oxygen. Several materials are capable of producing this effect. The most promising is $SrTiO_3$ anodes (74). Eutectics with $SrTiO_3$ may produce high conversion efficiencies by product or coupling effects.

CONCLUSIONS

Directionally solidified eutectic materials can have unique and interesting properties. These materials may be useful for electronic, magnetic or optical applications. The areas for future application are communications,

energy conversion and radiation detection.

The use of product effects will allow the design of a eutectic material with a specific property for a specific application. The design of the microstructure through controlled directional solidification may then be used to optimize the desired effect or to make the effect directional.

In the search for in-situ composite materials we should not be limited to only eutectics. Eutectoid and vitreous transformation reactions should also be considered.

Finally, for many applications the expense and fabrication difficulties of an in-situ composite may not be necessary. The artificial coupling of the two or more materials may provide the desired effect. When thinking of a material system for a given component or effect do not limit yourself. Think of simple single phase and two phase systems, artificial composites, as well as directionally transformed vitreous, eutectoid and eutectic materials.

REFERENCES

1. van Suchtelen, J., Philips Res. Repts. 27, pp 28-37, 1972.
2. Hansen, M., Constituion of Binary Alloys, pg 1107, 1958.
3. Fleming, M.C., Solidification Processing, pp 31-58, 1974.
4. Jackson, K.A., and Hunt, J.D., Trans. TMS-AIME, 236, pp 1129-1142, 1966.
5. Hunt, J.D., J. Inst. Metal. 94, pp 125-129, 1966.
6. Hunt, J.D., J. Cryst. Growth, 3,4, pp 82-91, 1968.
7. Day, M. and Hellawell, A.J., J. Austr. Inst. Metals, 9, pg 213, 1964.
8. Kerr, H.W., and Winegard, W.C., J. of Metals, 18, pp 563-569, 1966.
9. Kerr, H.W. and Winegard, W.C., Crystal Growth 1966, paper B25, pp 179-182, 1966.
10. Hogan, L.W., Kraft, R.W., and Lemkey, F.D., in Advances in Materials Research, Vol. 5, Ed. H. Herman, pp 83-211, 1971.
11. Thompson, E.R., and Lemkey, F.D., "Directionally Solidified Eutectic Superalloys," in Composite Materials, 4, Ed. Kreider, K.G., pp 101-157, 1974.

12. May, G.J., and Jones, D.R.H., Metals and Materials, 1974, pp 242-244.

13. Brady, F.L., J. Inst. Metals, 28, No. 2, pp 369-420, 1972.

14. Galasso, F.S., J. of Metals, 19, pp 17-21, 1967.

15. Galasso, F.S., in High Modulus Fibers and Composites, pp 87-112, 1969.

16. Salkind, M.J., Lemkey, F.D., and George, F.D., in Whisker Technology, Ed. Levitt, A., pp 343-401, 1970.

17. May, J. Metals and Materials, 1975, pp 44-48, 1975.

18. Salkind, M. and Lemkey, F., International Sci. and Tech., pg. 52, 1967.

19. Weiss, H., Met. Trans, 2, pp 1513-1521, 1971.

20. Goodrum, J.W., Georgia Institute of Technology, private communications.

21. Weiss, H., "Physical Properties of In Situ Composites" in Conference on In Situ Composites- II, September 2-5, 1975, pp 377.383.

22. Liebman, W.K. and Miller, E.A., J. Appl. Phys. 34, pp 2653-2659, 1963.

23. Weiss, H. and Wilhelm, M., Zeitschrift fur Physik, Band 176, pp 399-408, 1963.

24. Paul, P., Weiss, H., and Wilhelm, M., Solid State Electronics, Pergamon Press, Vol. 7, pp 835-841, 1964.

25. Weiss, H., Electron News, October 12, 1964, pg. 42.

26. Weiss, H., Solid State Electronics, Pergamon Press, Vol. 9, pp 443-451, 1966.

27. Elliott, C.T., and Hiscocks, S.E.R., Brit. J. Appl. Phys. (J. Phys. D.) Ser. 2, pp 1083-1087, 1969.

28. Umehara, Y., and Koda, S., Metallography, 7, pp 313-331, 1974.

29. Helbren, N.J., and Hiscocks, S.E.R., J. Mat. Sci., 8, pp 1744-1750, 1973.

30. Levinson, L.M., NMAB 308-III, pp 97-109, 1973.

31. Hellawell, A., "The Growth and Structure of Eutectic Silicon and Germanium," Progress in Materials Science, Vol. 15, Pergamon Press, pp 1-78.

32. Albers, W., and Verberkt, J., J. Mat. Sci., 5, pp 28-28, 1970.

33. Yue, A., U.C.L.A., private communications.

34. Favier, J.J. et al, In-Situ Composites 1975, pp 415-424.

35. Cullity, B.D., Introduction to Magnetic Materials, Addison-Wesley, 1972.

36. Tebble, R.S., and Crack, D.J., Magnetic Materials, Wiley-Interscience, 1969.

37. Albright, D.L., and Kraft, R.W., Trans. AIME, 236, pp 998-1003, 1966.

38. Albright, D.L., Ph.D. Dissertation, Lehigh University.

39. Galasso, F.S. et al., J. Appl. Phys. 38, pp 3241-3244, 1967.

40. Galasso, F.S. et al., J. Amer. Cer. Soc., 50, pp 333-334, 1967.

41. Jackson, M.R., Tauber, R.N. and Kraft, R.W., J. Appl. Phys. 39, pp 4452-4457, 1968.

42. Yim, W.M. and Stafko, E.J., J. Appl. Phys., 38, pp 5211-5216, 1967.

43. Collings, D.A. and Kossowsky, R., Met. Trans, 2, pp 1523-1527, 1971.

44. Livingston, J.D., J. Appl. Phys., 41, pp 197-204, 1970.

45. Sahm, P.R., and Killias, H.R., J. Mat. Sci., 5, pp 1027-1037, 1970.

46. Lotham, J.G., and Hellawell, A., J. Amer. Cer. Soc., 47, pp 184-188, 1964.

46a. Penfold, D., and Hellawell, A., J. Amer. Cer. Soc., 48, pp 133-135, 1965.

47. Batt, J.A., Douglas, F.C., and Galasso, F.S., Ceramic Bulletin, 48, pp 622-626, 1969.

48. Yue, A.S., and Yu, T.G., "Optical Properties of NaCl-NaF Eutectics," In-Situ Composites, 1975, pp 425-431.

49. Sievers, A.J., "Optical Properties of Composite Structures" in NMAB-308-III, January 1973.

50. Bright, M.W.A., and Lewis, M.H., J. Mat. Sci., 6, pp 1012-1020, 1971.

51. Bright, M.W.A., and Lewis, M.H., J. Mat. Sci. 6, pp 1246-1253, 1971.

52. van Hoof, L.A.H., and Albers, W., "Conversion of X-rays into Visible Light in NaCl-$PbCl_2$ Composite System," In Situ Composites - 1975, pp 407-413.

53. Davis, N.M., Clawson, A.R., and Weider, H.H., Appl. Phys. Letters, 15, pp 213-215, 1969.

54. Cline, H.E., Rose, R.M. and Wulff, J., J. Appl. Phys. 34, pg 1771, 1963.

55. Levy, S.A., Kim, Y.B., and Kraft, R.W., J. Appl. Phys. 37, pg 3659, 1966.

56. Gerdes, R.J., Chapman, A.T., and Clark, G.W., Science 167, pp 976-980, 1970.

57. Chapman, A.T., Gerdes, R.J., Wilson, J.C., and Clark, G.W., J. Cryst. Growth, 13/14, pp 765-771, 1972.

58. Chapman, A.T., Clark, G.W., and Hendrix, D.E., J. Amer. Cer. Soc., 53, pp 60-61, 1970.

59. Goodrum, J.W., Georgia Inst. of Technology, private communications, 1977.

60. Hulse, C.O., and Batt, J.A., Advanced Materials, Composites and Carbon, Sup. Amer. Cer. Soc., April 26-28, 1971, pp 133-141.

61. Hulse, L.O., and Batt, J.A., MClC, Rev. Cer. Tech. September 1974, Issue 33.

62. Schmid, F. and Viecknicki, D., J. Mat. Sci., 5, pp 470-473, 1970.

63. Livingston, J.D., J. Cryst. Growth, 24/25, pp 94-101, 1974.

64. Carpay, F.M.A., J. Cryst. Growth, 18, pp 124-128, 1973.

65. Carpay, F.M.A., Acta Met. 20, pp 929-933, 1972.

66. Carpay, F.M.A., and Cense, W.A., J. Cryst. Growth, 24/25, pp 551-554, 1974.

67. van der Boomgaard, J. et al., J. Mat. Sci., 9, pp 1705-1709, 1974.

68. Press release from Philips Research, 1974.

69. van der Boomgaard, J. and van Run A.M.J.G., In-Situ Composites 1975, pp 433-440.

70. Robertson, A.L., Research News Section, Science, 185, pp 846-847, 1974.

71. Quinlan, K., Deputy for Electronic Technology, Hanscom AFB, MA. private communication, 1977.

72. Sisson, R.D., Memo to C. S. Sahagian, Hanscom AFB, MA., 1977.

73. Yue, A.S., U.C.L.A., private communication, 1977.

74. Mavroides, J.G., Kafalas, J.A., and Kalesar, D.F., Appl. Phys. Letters, 28, pp 241-243, 1976.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by:
The Air Force Office of Scientific Research

Conducted by:
Auburn University and Ohio State University

PARTICIPANTS FINAL REPORT

SYNTHESIS AND CHARACTERIZATION OF NOVEL NITROALIPHATIC MATERIALS

Prepared by:	Melvin L. Druelinger, Ph.D.
Academic Rank:	Associate Professor
Department and University:	Department of Chemistry Indiana State University
Assignment:	
(Air Force Base):	USAF Academy
(Laboratory):	Frank J. Seiler Research Laboratory
(Directorate):	Chemical Sciences
(Division):	Energetic Materials Chemistry
USAF Research Colleague:	Capt Scott A. Shackelford, Ph.D.
Date:	August 19, 1977
Contract No.:	F44620-75-C-0031

SYNTHESIS AND CHARACTERIZATION OF NOVEL NITROALIPHATIC MATERIALS

By

Melvin L. Druelinger

ABSTRACT

New classes of energetic compounds for use in explosive fills, solid rocket propellants, detonation devices, etc., are of obvious interest. A very promising approach to achieving greater energy content as well as better thermal and mechanical stability is to develop energetic materials that could replace the nonenergetic binders, plasticizers, oxidizers, and crosslinking agents currently in use. The resulting improvement in performance characteristics of such materials would lead to increased range/payload capabilities, improved storage properties, and more efficient blast effects.

The value of the nitro and fluoro groups as constituents of energetic materials is well established as energetic and thermally stabilizing moieties, respectively. The introduction of an acetylenic group into a molecule can offer some additional advantages. The triple bond not only possesses significant energy content but is also more energetic than the analogous vinyl bond. The acetylenic moiety frequently appears to decrease a material's shock sensitivity and could increase the thermal stability of molecules in which it is present. Additionally, it can serve as the locus for further chemical modification of the molecule through facile addition and coupling reactions characteristic of this group. Finally, the triple bond itself, or an alkene easily derived from it, can be used for polymerization reactions. A molecule containing all three of these desireable molecular moieties, nitro, fluoro, and acetylene, would be of interest as a new energetic material.

This report describes a detailed study of a reaction designed to give such materials. Reaction variables and structural variations have been carefully examined. The reaction system involves the generation of a lithium acetylide salt, and its subsequent interaction with polynitro and fluoropolynitro triflates. The resulting products are conjugated enynes containing fluoro and nitro groups. Isolated yields as high as 84% have been obtained in favorable cases. Current data suggest wide variations in yield (0-84%) as both the nature of the acetylide and the triflate ester are varied. Experimental conditions also appear to effect yields significantly.

Present information suggests yields are best in the case of aryl acetylides reacting with 2-fluoro-2, 2- dinitroethyl triflate. The synthetic utility of this system and the scope of the reaction have been substantially defined, and demonstrated. Applications of these materials as energetic polymers for the variety of uses mentioned remain to be explored.

ACKNOWLEDGEMENTS

The author is grateful to the Air Force Systems Command for support of this research and for the opportunity to become acquainted with a number of areas of mutual interest. The summer experience has been greatly facilitated by the excellent administrative support provided through the ASEE and Auburn University by Mr. J. Fred O'Brien, and through the Air Force Academy by Lt Colonel Thomas M. Tomaskovic.

The personnel at the Frank J. Seiler Research Laboratory have been uniformly splendid in both professional and personal support. This unique combination is greatly appreciated by the author. Special thanks are due to Colonel M. D. Bacon and Lt Colonel Ben Loving for their attention, encouragement, advice, and hospitality.

It is with special pleasure that I acknowledge Captain Scott Shackelford, my research colleague, and Lt Robert Hildreth. As close working associates, they provided necessary assistance, advice, and support in an extremely professional and personable manner. Their enthusium and encouragement were invaluable.

It has been enjoyable and profitable working with the laboratory staff, and I wish to express my thanks to Mr. Lloyd Pflug and Mr. Fred Kibler for their excellent technical support, so openly extended in a cheerful manner.

I. INTRODUCTION

The Energetic Materials Chemistry Division of the Frank J. Seiler Research Laboratory (AFSC, USAF Academy) maintains an active research program that explores the preparation and properties of energetic materials that either are, or may be, used for explosive fills, solid rocket propellants, and related purposes. One of the more important goals of this program is to design and prepare new classes of energetic compounds that either have a greater energy content and/or superior stability (particularly thermal) than those currently in use. Another important goal is to develop energetic materials that could replace nonenergetic binders, plasticizers, oxidizers, and crosslinking agents currently used in the preparation of solid rocket propellants, and plastic bonded explosives.

The incorporation of energetic binder ingredients into solid propellant and explosive formulations offers an excellent approach for increasing the energy output during combustion and detonation processes, respectively. Currently nonenergetic binder materials are in wide use. While the formulated ingredients perform the required binding function, they are dead weight when compared to the energy output of the fuel/oxidizer or explosive materials used in solid propellants and munition charges. Therefore, these nonenergetic binder components add weight and consume space in propellant/munition cavities, while contributing no significant energy enhancement to performance results.

The substitution of energetic materials that perform binder functions could add significant energy contributions to combustion and detonation processes resulting in a higher specific impulse during propellant combustion and more efficient blast effects in explosives detonation. These increased performance characteristics would offer increased range or payload opportunities in strategic missiles, longer standoff ranges in tactical missiles, and higher yield detonation processes in munitions applications. Energetic binder development has suffered, however, from a lack of efficient high yield synthesis procedures suitable for production quantities. Such synthetic methods are required to produce energetic monomeric precursor compounds for subsequent polymerization into energetic binder materials that possess acceptable energetic, thermal, and hydrolytic stability properties.

It is well known that the introduction of one or more nitro groups will drastically increase the energy content of a molecule. The addition of a fluoro substituent usually lends a degree of thermal stability and increases the initial density of a molecule. Polyfluoro compounds often display critical lubricating properties. This is a potential plus for polymeric materials. With these concepts in mind, obvious target molecules for energetic monomers/polymers would contain polyfunctionality involving nitro and/or fluoro groups. The inclusion of an alkene linkage would permit a molecule to be easily polymerized and then used as an energetic plasticizer or binder component.

Compounds containing the acetylenic functional group as well as nitro and/or fluoro groups are of particular interest since the triple bond not only has considerable energy content, but is also more energetic than the analogous

vinyl bond. The acetylenic moiety frequently increases the thermal stability of molecules in which it is present. Additionally, triple bonded compounds are known, in some cases, to be less easily detonated by other munitions. It can also serve as the locus for further modification of the molecule through facile addition and coupling reactions characteristic of this chemical group. Finally, the triple bond itself, or an alkene easily derived from it, can be used for facile polymerization reactions.

A new reaction (described below) has been discovered in the F. J. Seiler Research Laboratory that produces a new class of energetic materials, namely nitroenynes. The products also incorporate both the alkene and acetylene functions and may include a fluoro substituent. These molecules produced by this reaction are of considerable interest since all of the structural elements delineated earlier are present in energetic monomers.

II. OBJECTIVES

The objectives of this work were to conduct a careful, detailed study of the scope and synthetic utility this new nitroenyne reaction might provide. This involved the evaluation of structural variations, reaction conditions, and work-up and isolation techniques. Such a detailed understanding of this reaction's scope would greatly increase its practical value in the production of new and useful energetic materials. It also would lead to increased yields and understanding of these and related substances. Any new compounds synthesized were to be structurally characterized. Finally, we planned to generate related new materials of interest by extending the concepts described to include alkoxides and simple alkyl carbanions as well as acetylides.

III. DISCUSSION AND RESULTS

Synthesis of New Nitroaliphatic Compounds Containing Acetylenic Groups

The most direct route to compounds of type III, featuring the desired functional groups, would be to carry out a nucleophilic displacement using an acetylide ion (I) and suitable derivatives (II) of the readily available highly nitrated alcohols (Z = OH),

$$R - C \equiv C^- + YC(NO_2)_2CH_2Z \rightarrow R - C \equiv C - CH_2C(NO_2)_2Y$$

I II III

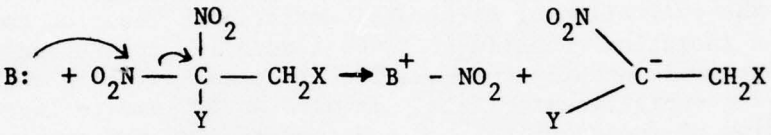
$$Y = (a)F(b)CH_3(c)NO_2$$

2-fluoro-2, 2-dinitroethanol (IIa), 2, 2-dinitro-1-propanol (IIb) and 2,2, 2-trinitroethanol (IIc). Unfortunately, two major problems hinder this approach. First, the starting alcohols have at least two large and very electronegative groups attached to the β-carbon. As a result, S_N2 displacements

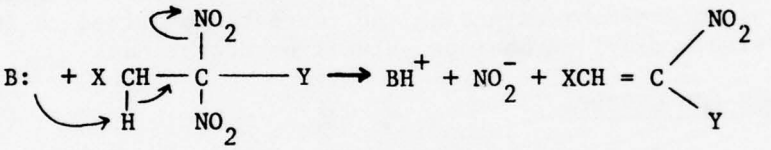
are seriously inhibited through both field and steric (neopentyl type) effects. This combination drastically inhibits similar reactions of α-halosulfones of type IV. [2] This lack of reactivity is further illustrated by the inertness of the benzenesulfonate derivatives of IIb toward acetate or chloride in refluxing ethanol. [3] Secondly, the nitro groups may interfere directly, either through direct attack of the nucleophile on one of the nitro groups affording a nitro-stabilized carbanion (eq. 1), or by inducing β-elimination of nitrite ion (eq. 2) [4]. Shackelford and Closson [1] were encouraged to try this approach, however, based on a recent report that described the triflate moiety as being an extremely labile leaving group in substitution reactions. [5] Additional encouragement was derived from the fact that α-substituted sulfones IVb [6] and IVc [7] undergo smooth substitution reactions with certain nucleophiles, and that some substitution reactions can be carried out on the tosylate IIa. [8]

RSO_2CH_2X X = (a)Cl, Br; (b)N_2^+(c)OSO_2CF_3

IV



(1)



(2)

Closson [1] chose to investigate the tosylate, bromide, and triflate derivatives of IIa which were available from literature preparations. The anticipated products III were not obtained. The most significant and surprising result was the discovery that acetylide ions react with β, β-dinitroalkyl trifluoromethanesulfonates (triflates, II, Z=OSO_2CF_3) to yield the unusual nitroenyne products. This signaled a new synthetic route to novel energetic materials that contain a number of significant chemical groups of interest. To summarize, polynitro and fluoropolynitro triflate esters can be prepared from the corresponding energetic alcohols and triflic anhydride (eq. 4). The latter is made from triflic acid (trifluoromethanesulfonic acid, eq. 3). Furthermore, these materials, when treated with lithium acetylides (from terminal alkynes and n-butyllithium), can be converted, surprisingly, to conjugated enynes possessing a terminal nitro (and/or fluoro) group on the alkene (eq. 5). Thus, these molecules and this reaction are of considerable interest since all of the structural elements delineated earlier as important to thermally stable energetic monomers are present in such molecules.

$$2\ TfOH + P_2O_5 \rightarrow TfOTf\ (Tf = CF_3SO_2^-) \qquad (3)$$

$$TfOTf + X-\overset{\overset{NO_2}{|}}{\underset{\underset{NO_2}{|}}{C}}-CH_2OH \xrightarrow[CH_2Cl_2]{pyr.} X-\overset{\overset{NO_2}{|}}{\underset{\underset{NO_2}{|}}{C}}-CH_2OTf \qquad (4)$$

V VI

X = (a)F(b)CH_3(c)NO_2

$$R-C\equiv C^-\ Li^+ + X-\overset{\overset{NO_2}{|}}{\underset{\underset{NO_2}{|}}{C}}-CH_2OTf \rightarrow R-C\equiv C-CH=C\begin{matrix}X\\ \diagdown\\ \diagup\\ NO_2\end{matrix} \qquad (5)$$

VII VI VIII

R = (a)ϕ(b)n - hex(c)EtO(d)CH_2 = C - CH_3(e)$(CH_3)_2$COLi

X = (a)F(b)CH_3(c)NO_2

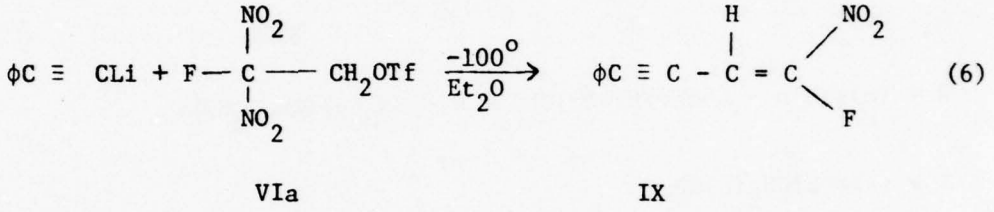
The purpose of the present research was to continue and extend the synthetic investigation of this reaction begun by Closson and Shackelford [1]. In particular, a careful and detailed study of the scope and synthetic utility of the process was needed. This necessitated work on the effects of structural variations, reaction conditions, work-up, and isolation techniques. Professor Closson's work, as directly related to this project, involved the acetylides VII a, c, and d, and the single triflate VIa. His best results were with VII (R = ϕ, X = F) where a 30% isolated yield was obtained. In other cases, yields were moderate and of an estimated, not isolated, nature. He reported that "isolation and purification presented considerable difficulty" and that distillation and sublimation usually destroyed most of the material.

With these results in mind, this summer's work was begun. All these triflates (VIa - c) were prepared and characterized. Improved yields were obtained in each case over those previously reported. The basic procedure involved mixing the energetic alcohol and freshly prepared triflic anhydride in methylene chloride at -23°C. A solution of pyridine in methylene chloride was added dropwise. The resulting mixture was filtered through silica gel

and chromatographed (yields were better than when distillation was used). The yields obtained were 63% for VIa, 85% for VIb, and 37% for VIc. The latter is a new compound and it was further characterized spectrosopically. It offers considerable potential for synthetic manipulations in this and related series of reactions. Some problems were encountered in its production, and yields can probably be substantially increased in future runs.

The requisite lithium acetylides were formed by adding n-butlyllithium to a solution of the terminal alkyne at 0^{o}, -23^{o}, or -78^{o}. Only in the case of lithium phenylacetylide was a reasonably homogeneous solution obtained, even at 0^{o}. In most cases a suspension was obtained. The product yields support the conclusion that high yields are best obtained from a homegeneous solution of the lithium salt. Future work must bear this carefully in mind.

The reaction studied most was that of lithium phenylacetylide and 2-fluoro-2, 2-dinitroethyl triflate (VIa). As shown in equation 6, the product is Z -1-fluoro-1-nitro-4-phenylbut-1-en-3-yne (IX). This previously observed transformation was verified in this study.



(6)

VIa IX

Of particular note is the fact that this material was obtained in 84% isolated yield (30% was the previous best [1]). This transformation clearly indicates that the new enyne forming reaction can, in favorable cases, be of considerable synthetic utility. The major experimental improvements permitting this remarkable success included inverse addition (organolithium to triflate), careful temperature control (-100^{o} to -78^{o} to -45^{o} to -23^{o} to $+1^{o}$) with TLC monitoring, avoidance of an aqueous work-up, and careful, even painstaking liquid chromatography using a UV monitor. The experiment modifications and improvements were used in all other reactions of this type.

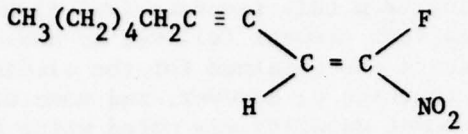
The importance of using inverse addition of the lithium salt as opposed to the "normal" addition (substrate added to precooled lithium salt) is dramatically illustrated by the following result. The reaction (eq. 6) was repeated identically in all respects, including work-up and isolation, except for the use of normal addition instead of inverse addition. The isolated yield dropped from 84% to 46%. Although this latter result is a modest yield, the reduction by nearly a factor of two is clearly synthetically significant.

All chromatographic fractions resulting from this reaction (eq. 6) and all others were carefully scrutinized by solvent removal followed by NMR spectroscopy of the residue. No firm evidence was obtained for the simple displacement product III. In the case of reaction 6, however, and some of the others, a reaction component of significant quantity was noted which had an R_f value greater than that of the enyne. In every case that intermediate vanished when the reaction temperature was raised to -45°C.

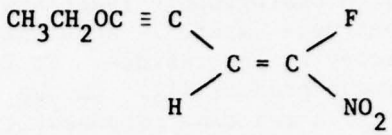
The number of other examples studied are indicated in equation 5. The reaction of lithium acetylides with triflate esters of β-nitro alcohols in diethyl ether at low temperatures to give conjugated nitroenynes like VIII appears to be quite general -- but limited as noted below. The mechanism by which they form is unknown, but a displacement/elimination process is favored. An attempt to ascertain this important information is underway. [9]

The enynes themselves are moderately stable yellow materials that appear to be readily polymerizable by acid or heat. None of the other examples studied came close to the yields obtained with the phenylacetylide system. When the nature of R is changed to alkyl (VIIb), the yield of isolated material (X) drops off drastically to 22%. This may be due in part to the fact that all alkyl systems possess a propargylic proton that is susceptible to removal by butyllithium. This complication does not arise with aryl systems. In addition, the alkyl group functions as an electron donor in contrast to the phenyl group in VIIa. The salt VIIb was very insoluble at 0° and reaction was not apparent even at -23°C. It was previously found [1] that the reaction of lithium phenylacetylide with VIa to produce IX proceeded best at temperatures below -78°C. At temperatures above -78°C little or no IX was found. The insolubility of VIIb at temperatures below 0°C necessitated conducting the reaction at this higher temperature and reasonably could be expected to contribute significantly to reduced enyne product yields. The problem of lithium acetylide insolubility at low temperature was encountered in all but the case of lithium phenylacetylide. Still, increased yields in this system may well be expected by reaction variable manipulation, but alkyl systems appear to be less useful than arylacetylenes.

The next best example resulted from the use of ethoxyacetylene. The yield in this system was about 13% of isolated material (XI). The low yield in this case may result from: 1) volatility of the product in the isolation process; 2) electron enhancement effect of the alkoxy group on the pi system; and 3) the use of "normal" addition due to the extreme volatility of the acetylene. Even at 0°, the lithium salt was very insoluble in ether. Once again this yield may be improved based on our present understanding of this reaction. The product was identical to that prepared by W. Closson [1] and R. Hildreth [10], and the yields are slightly better.



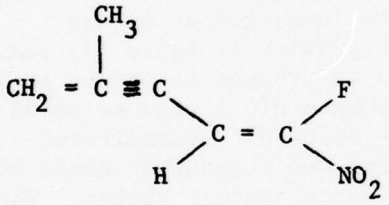
X



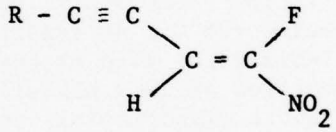
XI

The lithium salt of 2-methyl-1-buten-3-yne (VIId) gave even poorer results. The product XII was obtained in only a 3-7% yield. This result is a slight improvement over previous examinations of this system. [1] In this case, significant complications resulted when the lithium salt was found to polymerize spontaneously while being transfered. Avoidance of this problem may improve yields. The presence of an allylic (methyl) group adds a potentially complicating factor to this system.

In all, five cases of widely varying results have been reported. The presence of eneyne in each case is clearly distinguished in the NMR by the presence of a doublet at ~ 6.7δ with J_{HF} = 21 Hz. This coupling constant clearly indicates the trans or Z stereochemistry around the alkene bond [11] (XIII). No evidence was found for a cis or E stereochemical arrangement in any case.

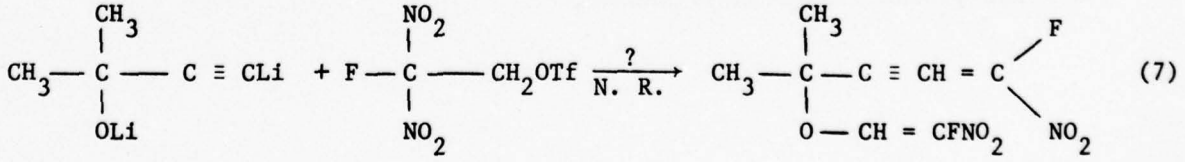


XII



XIII

In an effort to extend the scope of the reaction to include alkoxides and simple carbanions, reactions 7 and 8 were attempted. In the case of 7, where a bifunctional lithium salt was tried, no identifiable product was obtained. This could be the result of extreme salt insolubility in the reaction medium. This reaction should be reexamined because the reactant acetylene is a readily available, cheap material, the potential product is very interesting, and it may have considerable utility as can energetic monomer. The complete analysis of reaction 8 is still in progress, but current data are not encouraging.



(7)

$$CH_3CH_2CH_2CH_2Li + F-\underset{\displaystyle NO_2}{\overset{\displaystyle NO_2}{\overset{|}{\underset{|}{C}}}}-CH_2OTf \xrightarrow[N.R.]{?} CH_3CH_2CH_2CH_2CH=CFNO_2 \qquad (8)$$

The systems reported up to this point deal with structural variation in the nature of the acetylide (anion). Studies were also made varying the nature of the triflate (VIa - c). The results for VIc have been described in some detail. Using the best acetylide (VIIa) and the set of conditions which afforded the best yield of IX, the triflates VI b and c were examined. In the case of VIb (eq. 9), very little product was obtained, ($\leq$ 3% isolated). This disappointing result is in agreement with those of Closson [9] and Hildreth [10] under different conditions. It appears that the substitution of a methyl for a fluoro atom in the triflate leads to more side reactions and a more readily polymerizable product. Efforts with the triflate system VIb have been abandoned.

The trinitro triflate ester VIc was examined as a substrate for the first time. Special interest in this reaction results from the potential of generating a gem-dinitrovinyl system. Unfortunately, this reaction afforded no clearly identified product, although the NMR tantalizingly suggests up to 6% of a product that may be the desired one. Definitive confirmation is not available at this point. The reaction mixture darkens at -100° and considerable polymer is readily formed at room temperature in air.

In any event, from information available to date, it is clear that the esters VIb and c are much poorer substrates for this reaction than is VIa, the fluoro ester.

$$\phi - C \equiv CLi + CH_3-\underset{\displaystyle NO_2}{\overset{\displaystyle NO_2}{\overset{|}{\underset{|}{C}}}}-CH_2OTf \xrightarrow{?} \phi - C \equiv C-\underset{H}{\overset{CH_3}{C=C}}\underset{NO_2}{} \qquad (9)$$

VIb

$$\phi - C \equiv CLi + O_2N-\underset{\displaystyle NO_2}{\overset{\displaystyle NO_2}{\overset{|}{\underset{|}{C}}}}-CH_2OTf \xrightarrow{?} \phi - C \equiv C-\underset{H}{\overset{NO_2}{C=C}}\underset{NO_2}{} \qquad (10)$$

VIc

Experiments in progress (e.g., equation 8) will be completed and others are planned before the tenure of this summer is completed. Regretably they can not be included due to the time frame imposed by this report (e.g. work completed after 8-1/2 weeks). These key experiments include the reaction of 2-ethynylpyridine as another example of an aromatic acetylide. This material has been ordered for weeks and should arrive momentarily. Also included on the agenda is the reaction of lithium 2, 2, 2-trifluoroethoxide. Another attempt will be made to generate the dinitrovinyl system of equation 10 and efforts will be made to improve the yields in the alkyl system X.

TABLE I

REACTIONS OF LITHIUM ACETYLIDES WITH NITROTRIFLATES

Triflate	RC ≡ CLi	Mode of Addition	Present Yield (Isolated)
FDNEOTf(VIa)	ϕ	Inverse	84
FDNEOTf(VIa)	ϕ	Normal	46
FDNEOTf(VIa)	$CH_2 = C - CH_3$	Inverse	3-7
FDNEOTf(VIa)	$CH_3(CH_2)_5$	Inverse	22
FDNEOTf(VIa)	CH_3CH_2O	Normal	13
FDNEOTf(VIa)	$(CH_3)_2COLi$	Normal	0
FDNEOTf(VIa)	n-BuLi (Not Acetylide)	Inverse	0
TNEOTf (VIc)	ϕ	Inverse	$0 \leq 6$
DNPOTf(VIb)	ϕ	Inverse	$0 \leq 3$

IV. EXPERIMENTAL

Infrared (IR) spectra were determined on a Beckman IR-20 spectrometer; nuclear magnetic resonance (NMR) spectra were recorded on a Varian T-60A spectrometer. Chemical shifts are reported as ppm (δ) relative to tetramethylsilane; deuterochloroform was used as a solvent unless otherwise specified. Low resolution mass spectra (MS) were recorded on a DuPont 21-491 double focussing magnetic sector instrument or on a Hewlett-Packard 5990A GC/MS system. Separations were made on silica gel (60/200 mesh, J. T. Baker) using Isco model 328 fraction collector, an Isco UA-5 UV monitor, and a Pharmacia refractive index (RI) monitor.

Reactions were run under an atmosphere of dry nitrogen. Dry ether, used in all acetylide reactions, was dried over sodium and distilled immediately before use. Methylene chloride was reagent grade, distilled, and stored over molecular sieves. Hexane, analytical reagent grade, was distilled from sodium hydroxide and stored over molecular sieves. All reaction glassware was base-washed and oven dried before use. Thin layer chromatographic analyses were made using Eastman sheets, #2020, 100 μ; visualization was via UV.

As evidence of the ease of preparation of these compounds, a typical reaction will be described in some detail. Phenylacetylene (1.44g, 14.1 mmol) was weighed into a septum-sealed flask equipped with a magnetic stirrer and 15 ml of dry ethyl ether was added. The solution was cooled in an ice bath (0°), flushed with nitrogen, and 9.0 ml of 1.6 M n-butyllithium in hexane was added slowly via syringe. Stirring was continued for 45 minutes. The resulting slightly in cloudy solution was then transferred *via* syringe to a septum-capped dropping funnel that was also flushed with nitrogen. The dropping funnel was attached to a 100 ml flask containing 2-fluoro-2, 2-dinitroethyl triflate (2.00g, 7.0 mmoles) dissolved in 20 ml of dry ether. The system was flushed with nitrogen and sealed with a mineral oil bubbler to monitor the nitrogen flow during the course of the reaction. It was then cooled with stirring in a dry ice-diethyl ether bath (*ca.* -100°) and the phenylacetylide solution was added dropwise over one half hour. The solution rapidly turned a red-brown color; it was held at -100° for two hours, at -78° (dry ice-acetone) for one hour, at -45° (dry ice-chlorobenzene) for one hour and at -23° (dry ice-carbon tetrachloride) for one and a quarter hours. Analysis by TLC at regular intervals revealed an intermediate that vanished at -45°. The reaction solution was stored overnight at +1°. The reaction mixture (dark red-brown) was concentrated with chloroform on a rotary evaporator along with 2-2.5g of silica gel. This coated material was placed on a column containing 31g of 60/200 mesh silica gel. The solvent progression used was hexane, 1:1 hexane-methylene chloride, methylene chloride. Fractions were collected in 10-12 ml volumes. A UV monitor was used to detect the product which eluted in fractions 16-42, in hexane. Solvent removal afforded 1.12g (84%) of the desired enyne as a yellow solid. Spectral measurements were in agreement with those previously obtained by Closson and Shackelford. [1]

V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This report documents research involved in studying in some detail the scope and utility of a new reaction which produces a new class of energetic materials. A reasonable synthesis of nitrofluoroenynes has been achieved with yields ranging as high as 84%. These materials themselves probably will not be directly used as energetic materials, but they are excellent candidates as prepolymers for energetic binder/plasticizer use. The attachment of additional energetic groups through addition reactions should be quite feasible. The ready polymerizability of these compounds makes their use as starting materials for the development of energetic binders, plasticizers, or cross-linking agents appear very promising.

Structural variations in both the nature of the triflate and the lithium acetylide have been studied. Examples of an alkoxide and a simple alkyl carbanion have also been carried out. Variables such as temperature, mode of mixing reactants, and isolation techniques have been investigated. A favorable set of conditions has evolved (see EXPERIMENTAL), but uniformly satisfying results have not. The reaction is favored by high lithium salt solubility, low temperatures (-100°), inverse lithium acetylide addition, non-aqueous work-up, avoidance of distillation of the product, careful UV-monitored column chromatography, and the use of base-washed glassware. This work has been beset by problems relating to lithium salt insolubility, product volatility, ease of polymerization, and a multiplicity of side reactions.

The key factor seems to be, however, the nature of the reactants. Only the 2-fluoro-2, 2-dinitroethyl triflate (VIa) seems to afford good results among the triflates tried. More work should be done with the trinitroethyl triflate (VIc). Of the acetylides, only the lithium acetylide VIIa gives good results. It appears that aryl or at least electron-withdrawing acetylenes are required. Unfortunately, this research suggests the scope of the reaction is rather narrow.

These results suggest that further studies should be directed towards the use of other aryl acetylides and triflate VIa. In particular, a systematic study of p-substituted phenyl acetylides should be initiated. The use of nitro and polynitro substituted acetylides should be particularly rewarding for energetic materials. Other acetylides containing electron withdrawing substituents should be of interest. Finally, more information on the mechanism of the reaction is needed.

A comparison of reactions is given in TABLE I.

REFERENCES

1. (a) Unpublished observations, S. A. Shackelford and W. D. Closson, F. J. Seiler Research Laboratory, USAF Academy. (b) See also "Final Report for 1976 USAF-ASEE Summer Faculty Research Program" by W. D. Closson.

2. F. G. Bordwell and G. D. Cooper, J. Am. Chem. Soc., 73, 5184 (1951). F. G. Bordwell and W. T. Brannen, J. Am. Chem. Soc., 86, 4645 (1964).

3. L. W. Kissinger, T. M. Benziger, H. E. Ungnade, and R. K. Rohwer, J. Org. Chem., 28, 2491 (1963).

4. L. Zeldin and H. Schechter, J. Am. Chem. Soc., 79, 4708 (1957).

5. C. D. Beard, K. Baum, and V. Grakauskas, J. Org. Chem., 38, 3673 (1973).

6. J. B. F. N. Engberts and B. Zwanenburg, Tetrahedron, 24, 1737 (1968).

7. H. Hovius and J. B. F. N. Engberts, Tetrahedron Lett., 2477 (1972).

8. H. G. Adolph, J. Org. Chem., 36, 806 (1971).

9. Personal communication, Professor W. D. Closson, Department of Chemistry, State University of New York at Albany.

10. Personal communication, Lt. Robert Hildreth, F. J. Seiler Research Laboratory, USAF Academy, Colorado.

11. R. M. Silverstein and G. M. Bassler, "Spectrometric Identification of Organic Compounds," John Wiley and Sons, Inc., New York, 1967, p. 145.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

OPTIMAL WEAPON ASSIGNMENT IN A TIERED AIMPOINT SYSTEM

Prepared by:	Jerren Gould, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Mathematics Claremont Graduate School
Assignment:	
(Air Force Base)	Air Force Academy
(Laboratory)	Frank J. Seiler Research Laboratory
(Division)	Director of Aerospace-Mechanics Sciences
(Branch)	Applied Mathematics
USAF Research Colleague:	Capt Salvatore J. Monaco
Date:	August 19, 1977
Contract No.:	F44620-75-C-0031

OPTIMAL WEAPON ASSIGNMENT IN A TIERED AIMPOINT SYSTEM

by
Jerren Gould

ABSTRACT

The supply of nuclear weapons may be considered as a scarce resource. A model of the deployment and delivery of the nuclear weapons to targets within a tiered aimpoint system is proposed. The model is intended as an analytical tool to assess the present assignment of weapons and to evaluate policy alternatives concerning nuclear weapon deployment and targeting.

While there are numerous factors which pertain to the assignment problem in its actual military application, by focusing on the damage evaluation, supply, range, coverage, and MIRV aspects, a nonlinear integer programming formulation is given to address the question of obtaining optimal (in the sense of military objectives) solutions to the assignment problem. It has been found experimentally that many associated linear programming problems have integral optimal solutions. Hence, a branch-and-bound technique is suggested as a computationally feasible solution algorithm. Minimum "cost" formulations of the assignment problem are also given.

1.0 Introduction

This research project, pursued under the auspices of FJSRL and SASM, is concerned with the optimal assignment of (nuclear) weapons in a tiered aimpoint system. When one is unable to place a weapon on every installation to attain a minimum specified damage on each installation or a limited strike scenario is contemplated, the supply of weapons may be considered as a scarce resource which must be assigned to targets in such a way as to optimize the military objectives. On the other hand, when it is possible to achieve the prescribed military objectives, it may be advantageous to obtain the assignment of weapons which is optimal with respect to other criteria such as cost, reserves, and combat efficiency.

There are numerous factors and constraints which pertain to this assignment problem in its actual application. This report addresses the mathematical modeling and analysis of the assignment problem focusing on damage evaluation and supply, range, and coverage constraints. The problem of MIRV weapons systems is separately discussed.

The goal of this study is to derive a computationally feasible algorithm to obtain an optimal weapons assignment in our restricted problem with a view toward modification as additional factors and constraints are incorporated.

2.0 Statement of the Problem

There are three types of vehicles in which (nuclear) weapons may be delivered to a target: ICBM, SLBM, and manned bomber. Any individual weapons system may be categorized according to

a. launch site
b. range
c. number of weapons contained in the delivery vehicle
d. yield of each weapon
e. restrictions on the dispersion of the weapons in the single delivery vehicle.

Based on these characteristics, it may be determined whether a particular target is accessible to a given weapons system. There are several levels of weapon yields; these levels are called tiers. All of the weapons carried in a single delivery vehicle must be targeted in a neighborhood of the others. For ballistic missiles such a region is called a footprint; for manned bombers a reasonable flight path must be demonstrable.

The territory of a potential adversary contains numerous installations (factories, missile silos, airfields, etc.) and each installation has been allocated value points depending on its strategic/tactical significance. For each tier the collection of all installations has been partitioned into subsets such that for each subset

a. the aimpoint, or designated ground zero (DGZ), is given

b. the value extracted from each installation is known

c. there is no value extracted for installations not in the subset.

The extracted value is the point count which takes into account the value of the installation and the expected damage inflicted by a weapon from the associated tier. Value may be extracted from an installation by an appropriately targeted weapon from any tier and the extracted value may vary as the tier.

A measure of the effectiveness of any assignment of weapons to DGZs is the sum of the values extracted from each installation. An interesting nuance of this evaluation is that if an installation is covered more than once, then the values extracted from that installation are not additive in the usual sense.

The main treatment herein will be consideration of the case where there are insufficient numbers of weapons to extract value from every installation. This case may arise when there is an insufficient supply of weapons or when a limited strike is contemplated. Such a scenario is called a target-rich environment. We shall, however, formulate other important scenarios.

PROBLEM STATEMENT

Given the number and deployment of weapon systems, our objective is to establish a feasible computer algorithm to identify optimal (or nearly optimal) assignments of weapons to DGZs in a target-rich environment in such a way as to accommodate the dispersion constraints on weapons in each single delivery vehicle.

3.0 Mathematical Formulation

3.1 Launch Sites

Without any loss of generality we shall assume that the weapons systems at any launch site are identical. Thus, at any single launch site the delivery vehicles have the same range and carry the same number of weapons; also, each of the weapons carried in these delivery vehicles have the same tier. The ranges of all the weapons at any launch site are the same and, hence, the set of DGZs any weapon can access is common to all weapons originating from the same launch

site. If there are distinct weapons systems deployed at the same location, we shall consider these as distinct launch sites.

Accordingly, this leads to a large number of launch sites in which only a few weapons are deployed at each. We know, however, that ICBMs are deployed in launch fields. A launch field is a relatively small region which contains launch sites with a common type of delivery vehicle each carrying the same number of weapons of a common tier. Since the distance from these launch sites to the DGZs in range is very much larger than the distances between launch sites, the set of DGZs in range of any weapon is almost identical to the set of DGZs in range of any other weapon from the launch field. Hence, we reduce the magnitude of the problem by establishing a representative launch site for every launch field. The geographic coordinates of the representative launch site may be obtained by taking the average, weighted by the number of weapons at each launch site, of the geographic coordinates of all of the launch sites in the launch field. The representative launch site shall contain all of the delivery vehicles and weapons in the launch field. We can similarly find a representative launch site for SLBMs by considering the centroid of any patrol region.

Henceforth, we shall only consider representative launch sites. So at any such launch site, say ℓ, all of the weapons have the same tier; therefore, knowledge of ℓ implies knowledge of the tier of weapon therein deployed. Hence, there is a function t_1 which yields the common tier of weapon deployed at launch site ℓ, namely $t_1(\ell)$. Let n_ℓ denote the number of weapons deployed at launch site ℓ and let μ_ℓ be the number of weapons carried in each delivery vehicle at launch site ℓ. Clearly, n_ℓ/μ_ℓ is the number of delivery vehicles deployed at launch site ℓ. Let (θ_ℓ,ϕ_ℓ) be the longitude and latitude coordinates of launch site ℓ. Let R_ℓ be the range of the weapons originating from launch site ℓ. We shall denote the number of launch sites by L.

If launch site ℓ contains ballistic missiles and $\mu_\ell > 1$, then the weapons common to a single delivery vehicle are said to be MIRVed. Each of these weapons must be targeted in a neighborhood of the others. Such a region is called a footprint. The footprint constraint is identical for each delivery vehicle deployed at the same launch site. At present, the lowest tier (highest yield) weapons are not MIRVed. If launch site ℓ contains manned bombers and $\mu_\ell > 1$, the dispersion constraint necessitates that a reasonable flight path be demonstrable.

3.2 Tiered Aimpoint System

The tier of a weapon corresponds to its yield inversely. That is, the lowest tier corresponds to the highest yield and the highest tier corresponds to the lowest yield. We shall assume that the tier ranges over the integers 1 to T.

The territory of a potential adversary is divided into complexes (target islands). A complex is a region containing installations such that a hit by a weapon at any point within the region will not affect installations in any other complex.

Within any complex the DGZs for the largest yield (tier 1) of weapon are found initially. The first DGZ of tier 1 is placed at or near the installation having the largest value. The exact position of the DGZ depends upon determination (by simulation) and maximization of the total value extracted from the installations in the vicinity of the DGZ. Clearly, hard installations of high value require a nearby hit to extract value, whereas softer installations are [redacted] as sensitive to the position of the DGZ. These installations in the vicinity [redacted]e DGZ are said to be covered by the DGZ. Various other parameters, such as [redacted] of burst, are determined simultaneously with the location of the DGZ. If [redacted]f the installations [redacted] complex are not covered by the first DGZ, a second DGZ of tier 1 i[redacted]rly placed at or near the installation with the largest value not yet [redacted] by a tier 1 DGZ. This process continues until all installations in the c[redacted]plex are covered by a tier 1 DGZ.

Once the tier 1 DGZs have been determined for a complex, the DGZs for the higher tiers (lower yields) of weapons may be obtained. Consider the region containing the installations covered by any individual tier 1 DGZ as a subcomplex. Clearly, the tier 1 DGZs within any complex partition that complex into subcomplexes. Within each subcomplex the higher tier DGZs are found, tier by tier, in the same manner that the tier 1 DGZs are found within the complex. However, the number and layout of DGZs in any higher tier must be found so that the sum of the values extracted from the installations within the subcomplex is (most often) within 10% of the value extracted with the tier 1 weapon. Certain hardened installations require more than 1 weapon (all of the same tier) of higher tier to extract sufficient value. In such a case we consider this single hardened installation as several dummy installations. Each such dummy installation may be covered by a higher tier weapon without the others necessarily being covered. The values extracted for these dummy installations are constructed so

Fig. 1. An Example of a DGZ Layout with T = 3 and D = 29 and an Accessibility Table with L = 4

we may consider the additional value extracted for an additional weapon targeted to the same position. Thus, the set of installations covered by any tier 1 DGZ (subcomplex) is partitioned for each higher tier into DGZs of that higher tier.

The construction of the DGZs depends on the values given to the installations. This, in turn, depends on the strategic/tactical objectives. These objectives also determine the installations under consideration. Clearly, distinct military objectives may lead to distinct layouts of DGZs. We shall only consider a fixed layout of DGZs as determined by a single set of military objectives. Assignments of weapons to DGZs may be compared for distinct sets of military objectives.

We label all of the DGZs for single layout as $j = 1, \ldots, D$, where D is the number of DGZs. Each DGZ has a unique tier even though the coordinates of several aimpoints are identical. So there is a function t_2 which yields the associated tier of the DGZ. That is, $t_2(j)$ is the associated tier of DGZ j. Let $I(j)$ be the set of installations covered by DGZ j. By construction, for each tier $t = 1, \ldots, T$, $\{I(j) | j \in t_2^{-1}(t)\}$ is a partition of the set of all installations. So clearly, if $I(j_1) \bigcap I(j_2) \neq \emptyset$ and $t_2(j_1) = t_2(j_2)$, then $j_1 = j_2$.

A special type of layout of DGZs is the completely nested layout. A layout of DGZs is completely nested if $I(j_1) \bigcap I(j_2) \neq \emptyset$ implies $I(j_2) \subset I(j_1)$ whenever $t_2(j_2) \geq t_2(j_1)$. In a completely nested layout, the installations covered by any higher tiered DGZ are contained in only one DGZ of each lower tier. Alternatively, the DGZs of tier $t+1$ which share installations in common with a tier t DGZ actually partition this set of installations. By construction, the higher tier DGZs are nested within tier 1, but the layout is not necessarily completely nested. As we shall see later, the completely nested case maintains the mathematical structure of the problem, but is computationally smaller problem; hence, the completely nested case is useful for testing computational algorithms in reducing computation time.

It is important to note that the layout of DGZs is constructed without regard for the weapons deployment and ranges. If there is no assignment of weapons to DGZ so that every installation is covered, we have a target-rich environment. Otherwise, we have a target-poor environment. We forego discussion of the target-poor case until § 5.0.

3.3 Range

Each DGZ is constructed to accept a weapon of only one particular tier. That is, only a weapon of tier $t_2(j)$ may be assigned to DGZ j. We say launch

site ℓ matches DGZ j if the weapons deployed at launch site ℓ are of the same tier as that required by DGZ j. Clearly, launch site ℓ matches DGZ j if and only if $t_1(\ell) = t_2(j)$. It is a necessary (but not sufficient) condition for a weapon to be assigned from launch site ℓ to DGZ j that launch site ℓ matches DGZ j.

Even if launch site ℓ matches DGZ j, a weapon may not be assigned from launch site ℓ to DGZ j unless DGZ j is in range of launch site ℓ. Thus, the distance from launch site ℓ to DGZ must be less than R_ℓ, the range of weapons originating from launch site ℓ. We shall assume that the ground trace of the delivery system and the reentry vehicle is the shortest distance along the surface of the earth between the launch site A and the aimpoint B; the ground trace is an arc of a great circle. We may use spherical trigonometry to resolve the distances and directions over the surface of the earth.

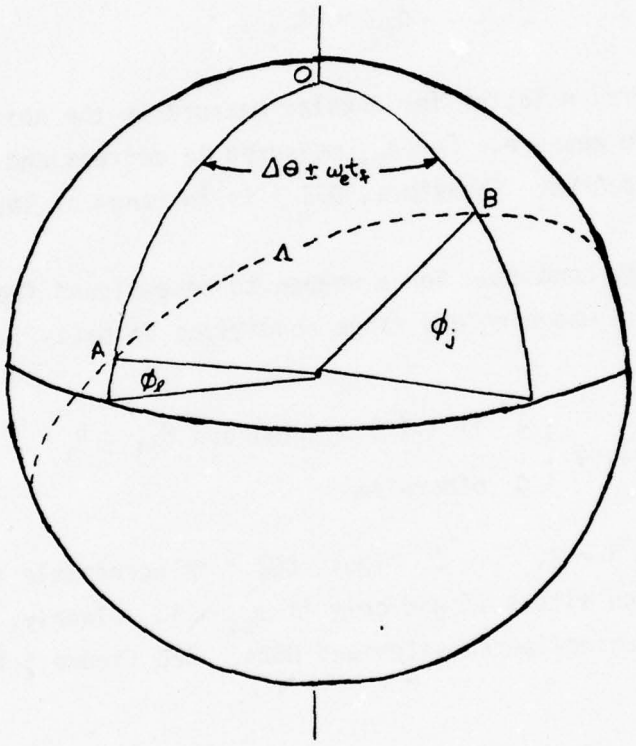


Fig. 2. Launch Site and Aimpoint at the Instant of Launch

Point O represents the north pole. Point A represents the launch site; its longitude and latitude coordinates are (θ_ℓ, ϕ_ℓ). Point B represents the aimpoint; its longitude and latitude coordinates are (θ_j, ϕ_j). Let ω_e denote the angular velocity of the earth and t_f denote the time of flight of the weapon. Hence, $\omega_e t_f$ is the earth rotation correction angle during the flight of the weapon. A

positive sign for $\omega_e t_f$ denotes an eastward launch whereas a negative sign denotes a westward launch. Let $\Delta\theta$ denote the difference in launch site and aimpoint longitudes.

OAB represents a spherical triangle whose two sides OA and AB and the included angle $\Delta\theta \pm \omega_e t_f$ are known. AB is the total range angle $\Lambda_{\ell j}$ and is the side opposite the known angle. From the Law of Cosines for a spherical triangle we obtain the total range angle

$$\Lambda_{\ell j} = \cos^{-1} \left\{ \sin \phi_\ell \sin \phi_j + \cos \phi_\ell \cos \phi_j \cos (\Delta\theta \pm \omega_e t_f) \right\}. \qquad (3.1)$$

Arc AB is the ground trace of the path of the weapon; its length is calculated as

$$d_{\ell j} = \beta \Lambda_{\ell j} , \qquad (3.2)$$

where β is the conversion factor for angular measure on the surface of the earth to the usual distance measure. For $\Lambda_{\ell j}$ measured in degrees and $d_{\ell j}$ measured in miles, $\beta = 69.04$ mi/degree. Therefore, DGZ j is in range of launch site whenever $d_{\ell j} \leq R_\ell$.

It is a necessary condition for a weapon to be assigned from launch site ℓ to DGZ j that both the matching and range conditions be satisfied. We define the accessibility index

$$\alpha_{\ell j} \equiv \begin{cases} 1 & \text{if } t_1(\ell) = t_2(j) \text{ and } d_{\ell j} \leq R_\ell \\ 0 & \text{otherwise} \end{cases}$$

for $\ell = 1, \ldots, L$ and $j = 1, \ldots, D$. Thus, DGZ j is accessible to a weapon originating from launch site ℓ if and only if $\alpha_{\ell j} = 1$. Clearly, $d_{\ell j}$ need only be calculated for matching launch sites and DGZs. See figure 1 for an accessibility table.

3.4 Value

If a weapon is assigned to DGZ j, the tier of the weapon must match the tier of the DGZ. By construction, the value extracted from each installation covered by DGZ j is known; let v_{ji} be the value extracted from installation i when a weapon is assigned to DGZ j. These values extracted depend on the value of the installation, its hardness, and geographical relationship to the aimpoint. Stochastic variations, such as aiming uncertainties, have been incorporated by averaging. It is possible that damage may be incurred to installations not

covered by the DGZ, but we shall assume that, in terms value extracted, this damage is negligible.

Let

$$a_{\ell j} = \begin{cases} 1 & \text{if a weapon from launch site } \ell \text{ is assigned to DGZ } j \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, the decision variables $a_{\ell j}$ are the ones we must determine for a solution of the problem.

The values extracted from distinct installations are assumed to be additive. However, if weapons are assigned to DGZs which cover the same installation, the total value extracted from that installation should not be the sum of the values extracted from that installation. Let $J(i) = \{j | i \in I(j)\}$, be the set of DGZs which cover installation i. By construction, there is exactly one DGZ from each tier belonging to J(i). So,

$$\sum_{\ell, j \in J(i)} a_{\ell j} \, v_{ji} \tag{3.3}$$

is not a practical measure of the value extracted from installation i if more than one of the $a_{\ell j}$ is 1. A reasonable measure of the value extracted from installation i is the greatest value extracted from the installation for any weapon assigned to a DGZ which covers installation i; that is, since the v_{ji} are generally nonnegative,

$$\max_{\ell, j \in J(i)} a_{\ell j} \, v_{ji}. \tag{3.4}$$

Accordingly, in a target-rich environment we wish to find the assignment scheme which maximizes the sum of the values extracted for each installation. That is, we wish to find $\{a_{\ell j}\}$ to maximize

$$\sum_i \max_{\ell, j \in J(i)} a_{\ell j} \, v_{ji}.$$

Thus, only assignments which conduce a largest value extracted for some installation contribute to the objective; hence, others may be viewed as waste and should be reassigned elsewhere. Notice that the creation of several dummy installations for certain targets allows some value extracted to accrue when the target is hit more than once.

Due to certain strategic/tactical considerations, some installations must not be damaged. For each such "flagged" installation i we make v_{ji}, for each $j \in J(i)$, very large in the negative sense to insure that any reasonable assignment would deny a weapon to any DGZ covering installation i. However, for any installation which is not "flagged" it remains that $v_{ji} \geq 0$.

3.5 The Mathematical Model

Based on the preceding discussion, our objective is to find the assignment scheme $\{a_{\ell j}\}$ which maximizes the total value extracted subject to accessibility, deployment, and dispersion constraints. In mathematical terms, we have a non-linear integer programming problem in which we must find $\{a_{\ell j}\}$ to

$$\text{maximize} \quad \sum_i \max_{\ell, j \in J(i)} a_{\ell j} v_{ji} \qquad (3.5\ 0)$$

$$\text{subject to} \quad a_{\ell j} \leq \alpha_{\ell j} \qquad \forall \ell, j \qquad (3.5\ A)$$

$$\sum_j a_{\ell j} \leq n_\ell \qquad \forall \ell \qquad (3.5\ S)$$

$$a_{\ell j} = 0 \text{ or } 1 \qquad \forall \ell, j \qquad (3.5\ I)$$

$$\text{the above assignment satisfies the dispersion restrictions} \qquad (3.5\ D)$$

Although this mathematical statement of the problem appears simple, it includes, albeit subtlely, all of the structure and form of the solution for the factors we are presently considering.

The assignment $a_{\ell j} = 0 \quad \forall \ell, j$ is a feasible, but not necessarily optimal solution of (3.5 OASID). There are only a finite number of feasible solutions of (3.5 OASID). Hence, by simple enumeration an optimal (not necessarily unique) solution must exist. However, because of the computational magnitude of the enumeration with evaluation of (3.5 D) in any problem approaching the size of any actual application, the problem cannot be performed on electronic computation equipment with this algorithm. Because of the mathematical complexity and this computational difficulty, we separate the problem into two components, (3.5 OASI) and (3.5 D). That is, we solve (3.5 OASI) and then attempt to evaluate whether a solution or a slight perturbation of a solution also satisfies (3.5 D).

Again, the assignment $a_{\ell j} = 0 \ \forall \ell, j$ is a feasible, but not necessarily optimal solution of (3.5 OASI) and there are only a finite number of solutions of (3.5 OASI). Hence, by simple enumeration an optimal (not necessarily unique) solution of (3.5 OASI) must exist. A longest path algorithm would identify such optimal solutions, but again the magnitude of the computation for a real application renders this algorithm unimplementable on a computer.

4.0 An Algorithm for Solution of the Assignment Problem

We decompose problem (3.5) into an assignment problem (3.5 OASI) and a dispersion problem (3.5 D). The enumeration algorithm is not a workable method for any real application; hence, we shall make use of some special structure of the problem to derive an algorithm which may be applied to a real scenario on the computer.

4.1 A Simplification of the Model

The difficulty in computing the optimal solutions of problem (3.5 OASI) lies in the nonlinearity of the objective function (3.5 O). This nonlinearity is the major factor which necessitates the enumeration (implicit or explicit) procedure to obtain a solution. We noted in § 3.4 that only assignments which lead to a largest value for some installation contribute to the objective. In any optimal assignment every DGZ which receives a weapon must contain an installation in which it extracts the dominant value (see expression (3.4)). Thus, when an installation is hit more than once, some weapons effectively extract no value from the installation. Although it is possible for this case to occur in an optimal solution, in a target-rich environment it should happen rarely. That is, almost all assignments will be to DGZs in which there is no competition among the weapons to extract value from the installations covered. Recall from the partitioning that an installation may be covered more than once only by DGZs (and weapons) of distinct tiers. If the value extracted from any installation remains stable as the covering DGZ is varied (that is, there is little variation in v_{ji} for $j \in J(i)$), the gain in total value extracted for any assignment scheme $\{a_{\ell j}\}$ is marginal. Hence, we shall exclude any assignment scheme in which an installation is covered by more than 1 DGZ receiving weapons. This constraint is

$$\sum_{\ell, j \in J(i)} a_{\ell j} \leq 1 \qquad \forall i. \tag{4.1}$$

It is interesting to notice that expressions (3.3) and (3.4) are equal when constraints (4.1) and (3.5 I) are satisfied. Also, the objective function (3.5 O) transforms as follows:

$$\sum_i \max_{\ell, j \in J(i)} a_{\ell j} v_{ji} = \sum_i \sum_{\ell, j \in J(i)} a_{\ell j} v_{ji}$$

$$= \sum_\ell \sum_j a_{\ell j} \sum_{i \in I(j)} v_{ji}$$

$$= \sum_{\ell, j} a_{\ell j} v_j, \qquad (4.2)$$

where $v_j = \sum_{i \in I(j)} v_{ji}$ is the value extracted from all of the installations in DGZ j when an appropriately tiered weapon is so assigned. Examples of v_j are given in figure 1.

On the basis of the mild simplifying assumption (4.1), the assignment problem may be restated as

$$\text{maximize} \sum_{\ell, j} a_{\ell j} v_j \qquad (4.3\ O)$$

$$\text{subject to} \quad a_{\ell j} \leq \alpha_{\ell j} \qquad \forall \ell, j \qquad (4.3\ A)$$

$$\sum_j a_{\ell j} \leq n_\ell \qquad \forall \ell \qquad (4.3\ S)$$

$$\sum_{\ell, j \in J(i)} a_{\ell j} \leq 1 \qquad \forall i \qquad (4.3\ C)$$

$$a_{\ell j} = 0 \text{ or } 1 \qquad \forall \ell, j. \qquad (4.3\ I)$$

The advantage of this formulation is that the objective function (4.3 O) is a linear function of the decision variables. Notice also that problem (4.3 OASC) is a linear programming formulation. If the optimal solution to (4.3 OASC) is integral, then it will be an optimal solution to (4.3 OASCI). In fact, the polytope defined by (4.3 ASC) is rich in integral vertices. Hence, an algorithm using the branch and bound method with linear programming as the evaluation and fathoming device should converge rather quickly to an optimal solution.

4.2 Branch-and-Bound

Although the branch-and-bound technique for solution of integer programming is essentially enumerative, the use of linear programming methods will allow us to skip over many feasible combinations of the variables to converge quickly to an optimal integral solution. Moreover, an efficient linear programming code may be constructed utilizing the special structure of our problem.

The fundamental strategy of the enumerative approach is to develop a tree hierarchy of nodes representing candidate problems whose solution may provide a solution to the integer programming problem. The search adds nodes to the tree until a solution is found which is shown to be the best over all others which have been found or which may be found as a result of searching the remaining candidate problems. The basic components of an effective enumeration are:

a. scheme for recording the state of nodes considered and keeping track of those to be considered
b. means for identifying a solution when it has been found
c. rule to choose the next candidate problem for consideration
d. strategy for directing the search of the tree.

The branch-and-bound technique is a method of directing the tree search.

The strategy at a node is to solve the associated linear programming problem, the integer programming problem with the integrality constraints relaxed. If the solution is all integer, then it becomes a candidate solution. Otherwise, one of the nonintegral variables, say $a_{\ell j}$, is chosen and two new nodes are created in the tree with the additional restriction

$$(1)\ a_{\ell j} = 0 \quad \text{or} \quad (2)\ a_{\ell j} = 1.$$

If a candidate solution is found, it is compared to the best prior candidate solution, the incumbent solution. If it is better (greater), it replaces the incumbent. That branch of the tree is terminated since any further restriction of this problem can only result in suboptimal solutions.

If the solution at a node is not strictly integral but the objective value is not better (greater) than the objective value of the incumbent solution, then that branch may be terminated, or fathomed, since any further restriction of this problem can only result in suboptimal solutions.

The tree in our problem begins with only one node, namely problem (4.3 OASC). By (4.3 C) an integral solution at any node must also satisfy (4.3 I). Zero may

be used as the initial incumbent objective value corresponding to the feasible integral solution $a_{\ell j} = 0 \quad \forall \ell, j$.

Algorithms to execute the branch-and-bound technique may be found in [2], [8], and many other places. A seemingly efficient rule to choose the next candidate problem for solution in our problem is to consider the newest node created. This allows a new constraint to be added to the existing computational tableau; the dual simplex algorithm may be employed to quickly find the next solution without undue storage transfer. The computational magnitude of the problem necessitates a computer code which minimizes the amount of computer storage required; hence, the computer code should be specially designed to exploit the particular structure of our problem.

4.3 Flow with Gains

Problem (4.3 OASCI) may be equivalently formulated in the context of flow with gains. We, however, omit the proof of the equivalence.

For all j such that $t_2(j) < T$ define $G(j) = \{k | I(k) \cap I(j) \neq \emptyset \text{ and } t_2(k) = T\}$. G(j) is just the set of highest tier DGZs which have an installation in common with DGZ j. Let $N_j = \#G(j)$ be the cardinality of G(j). In the flow formulation, weapons flow from the launch sites to the DGZs. If a weapon is assigned to a lower tier DGZ, it is multiplied and proceeds to every highest tier DGZ which shares an installation in common with the DGZ to which the weapon was assigned. Thus, each DGZ of the highest tier may not receive more than one weapon, either directly or indirectly. Mathematically, the equivalent flow with gains formulation is

$$\text{maximize} \sum_{\ell,j} v_j \, a_{\ell j} \qquad (4.4\ 0)$$

$$a_{\ell j} \leq \alpha_{\ell j} \qquad \forall \ell, j \qquad (4.4\ A)$$

$$\sum_j a_{\ell j} \leq n_\ell \qquad \forall \ell \qquad (4.4\ S)$$

$$-N_j \sum_\ell a_{\ell j} + \sum_{k \in G(j)} b_{jk} = 0 \qquad \forall j \ni t_2(j) < T \qquad (4.4\ T)$$

$$a_{\ell j} = 0 \text{ or } 1 \qquad \forall \ell, j \qquad (4.4\ I)$$

$$b_{jk} = 0 \text{ or } 1 \qquad \forall j \ni t_2(j) < T \text{ and } \forall k \ni t_2(k) = T. \qquad (4.4\ I')$$

The variables $a_{\ell j}$ have the same interpretation as before. The variables b_{jk} represent the flow from lower tier DGZs to the highest tier DGZ. Notice the variables b_{jk} do not appear in the objective function.

The flow with gains formulation is presented here because of the possibility that it may lead to distinct, more efficient algorithms to solve the assignment problem.

4.4 An Example

The optimal solution to the problem presented in figure 1 is

$$a_{\ell j} = \begin{cases} 1 & \text{for } (\ell,j) \in \{(1,3),(2,7),(2,8),(3,14), \\ & \qquad (3,16),(4,27),(4,28)\} \\ 0 & \text{otherwise.} \end{cases}$$

5.0 Target-poor Environment

In a target-poor environment there are enough weapons to extract value from every installation we desire to damage; an optimal assignment according to military objectives, indeed, would include the damaging of all such installations. However, there may be many such assignments which achieve approximately the same military objective value. Thus, we may use other criteria to decide which of these optimal military assignments is best.

A crude measure of the density of targets is

$$P = \sum_{t=1}^{T} \frac{W_t}{D_t},$$

where W_t is number of weapons of tier t and D_t is the number of DGZs in tier t. If $P \leq 1$, we have a target-rich environment. Otherwise, we have a target-poor environment. P is an inexact measure of target-richness because it ignores the range constraints.

The target-poor assignment problem is formulated as

$$\text{minimize} \sum_{\ell,j} c_\ell \, a_{\ell j} \tag{5.1 O}$$

$$\text{subject to} \quad a_{\ell j} \leq \alpha_{\ell j} \qquad \forall \ell,j \tag{5.1 A}$$

$$\sum_j a_{\ell j} \leq n_\ell \qquad \forall \ell \tag{5.1 S}$$

$$\sum_{\ell,j\in J(i)} a_{\ell j} \geq 1 \qquad \forall i \tag{5.1 P}$$

$$a_{\ell j} = 0 \text{ or } 1 \qquad \forall \ell, j. \tag{5.1 I}$$

The variable c_ℓ is interpreted as the "cost" of a weapon from launch site ℓ. This designation of "cost" is very general and encompasses many distinct situations. If $c_\ell = 1 \ \forall \ell$, our objective would be to minimize the number of weapons to obtain the military objective. If $\mu_\ell c_\ell$ is the throwweight of a delivery vehicle deployed at launch site ℓ, we may use (5.1 OASPI) to minimize the total throwweight. We may also view c_ℓ as the yield of a weapon of tier $t_1(\ell)$ in order to minimize nuclear contamination. Also, $\mu_\ell c_\ell$ may be viewed as the monetary cost of a weapons system deployed at launch site ℓ; thus, we may obtain the most cost-effective way to achieve the military objectives. Notice that formulation (5.1 OASPI) does not depend on the rather subjective determination of values extracted. Once an installation is included in the layout, a weapon will be allocated to cover it. DGZs which contain "flagged" installations must be removed from the layout.

If we replace constraint (5.1 P) by

$$\sum_{\ell,j} v_j \, a_{\ell j} \geq V, \tag{5.1 V}$$

where V is military objective as expressed in value points extracted and add constraint (4.3 C), then problem (5.1 OASVI + 4.3 C) is a generalization of the original target-poor assignment problem. An interesting application of (5.1 OASVI + 4.3 C) in a target-rich environment is to use (4.3 OASCI) to find V and then solve (5.1 OASVI + 4.3 C). This would yield the least "cost," best military assignment in a target-rich environment.

Since the form of problem (5.1) is very similar to the form of problem (4.3), a similar branch-and-bound method may be used to obtain the optimal solution.

6.0 The Dispersion Problem

The dispersion problem arises due to the fact that the delivery vehicles deployed at some of the launch sites transport more than 1 weapon. These weapons cannot be assigned arbitrarily; each must be assigned to a DGZ near to the other DGZs which receive a weapon from the same delivery vehicle.

We shall again, without any loss of generality, employ the convention that the weapons systems deployed at any single launch site are identical. Thus, each delivery vehicle at a given launch site carries the same number of weapons of a common tier and the dispersion constraints are identical for the weapons carried on any such delivery vehicle. Notice that if $\mu_\ell = 1$, then there is no dispersion problem; hence, we shall only consider launch sites in which $\mu_\ell > 1$.

We shall assume that a solution to the assignment problem is given. Our objective is to determine whether such an assignment satisfies the dispersion constraints. There may be many optimal solutions to (4.3 OASCI); each must be checked to see if it satisfies the dispersion constraints. If no optimal assignment satisfies the dispersion constraints, methods must be developed to identify nearly optimal solutions which do satisfy the dispersion constraints.

In a target-rich environment all of the weapons allocated to a strike will be used. Hence, we may presume that all of the weapons contained in each delivery vehicle allocated to the strike have been assigned to a DGZ. For the target-poor environment case, this is an assumption.

6.1 MIRV

A ballistic missile which carries more than 1 weapon is said to be MIRVed. The weapons are carried in the delivery vehicle until it is near the target region; the weapons then separate from the delivery vehicle to proceed to the particular aimpoint each is assigned. The region into which these weapons must be aimed is called the footprint. The footprint is characterized by the weapon system. Hence, all weapons systems deployed at the same launch site have the same footprint.

We assume that the boundary of the footprint may be suitably approximated by a fixed ellipse. That is, the boundary of the footprint for each weapons system originating from launch site is defined by a major axis a_ℓ and a minor axis b_ℓ. The orientation of the major axis is along the great circle containing the launch site and the centroid of the DGZs assigned to that delivery vehicle and covered by the ellipse.

We decompose the dispersion problem for MIRVs by considering each launch site separately. When several assigned DGZs of the same tier are in range of several launch sites, then some permutations of the launch sites for these assignments will yield the same objective values. Thus, some DGZs may be traded among launch sites. Each such permutation must be examined. We shall check the feasibility of each assignment individually; hence, we assume a fixed assignment.

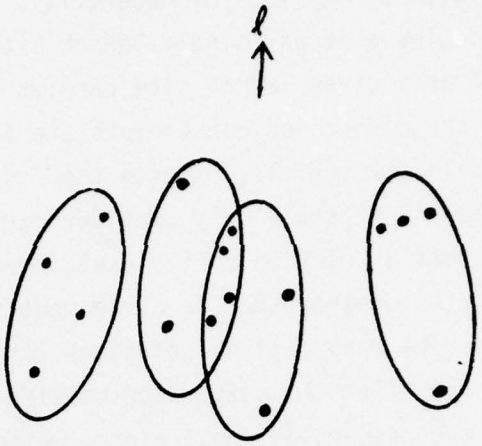


Fig. 3. An arrangement of assigned DGZs satisfying the dispersion constraint, $n_\ell = 16$, $\mu_\ell = 4$

The launch site ℓ is not a DGZ and the range of any weapon is less than half the circumference of the earth. Thus, if we view the launch site ℓ as a pole of a spherical earth, neither pole is an aimpoint. By using a Mercator map of the earth with the launch site as a pole, we may align the elliptical footprints so they are parallel. By a transformation of scale in horizontal direction of the Mercator map, we may transform the boundary of the footprint into a circle. Notice that the horizontal expansion depends on the vertical coordinate (distance from launch site). This transformation may even be generalized to allow original footprint shapes other than the ellipse. These transformations of the globe require that the coordinates of the DGZs receiving weapons originating from launch site ℓ be accordingly transformed so that dispersion feasibility may be attempted.

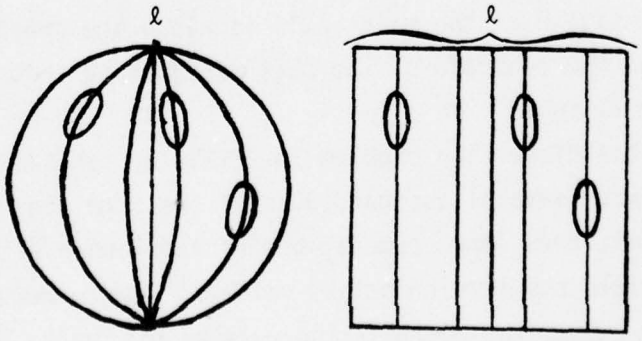


Fig. 4. The Mercator transformation

Since we are only considering one launch site at present, we suppress ℓ in our notation.

Let $s = n/\mu$ be the number of delivery systems deployed at launch site ℓ. Let ρ be the radius of the transformed footprint (circle). Let (x_j, y_j) be the transformed coordinates of DGZ j, which has been assigned a weapon from launch site ℓ, for $j = 1, \ldots, n$. We wish to determine (X_k, Y_k), $k = 1, \ldots, s$, the transformed coordinates of the center of the footprint for delivery system k in a MIRVed assignment. Let

$$\lambda_{jk} = \begin{cases} 1 & \text{if DGZ j receives a weapon from delivery vehicle k} \\ 0 & \text{otherwise.} \end{cases}$$

Consider the problem

$$\text{minimize } r^2 \tag{6.1}$$

$$\text{subject } \sum_j \lambda_{jk} = \mu \qquad \forall k$$

$$\sum_k \lambda_{jk} = 1 \qquad \forall j$$

$$\lambda_{jk}\left[(x_j - X_k)^2 + (y_j - Y_k)^2\right] \leq r^2 \qquad \forall j,k$$

$$\lambda_{jk} = 0 \text{ or } 1 \qquad \forall j,k.$$

If the optimal objective value of problem (6.1) does not exceed ρ^2, then the assignment from launch site may be MIRVed. The solution of problem (6.1) yields a yes or no answer to the question. On the other hand, a solution of

$$\text{maximize } \sum_k \gamma_k \tag{6.2}$$

$$\text{subject to } \gamma_k\left(\mu - \sum_j \lambda_{jk}\right) = 0 \qquad \forall k$$

$$\sum_k \lambda_{jk} = 1 \qquad \forall j$$

$$\gamma_k\left\{\rho^2 - \lambda_{jk}\left[(x_j - X_k)^2 - (y_j - Y_k)^2\right]\right\} \geq 0$$

$$\lambda_{jk} = 0 \text{ or } 1 \qquad \forall j,k$$

$$\gamma_k = 0 \text{ or } 1 \qquad \forall k$$

may be more enlightening. $\sum_k \gamma_k$ may be interpreted as the maximal number of MIRVs that can be used from launch site under the present assignment. Solution of (6.2) also yields their construction.

Weaker notions of MIRV compatibility may be derived by consideration of the Weber problem and Cooper's algorithm ([1],[6]). In such a formulation the requirement that a delivery system carry a fixed number of weapons is relaxed and the objective is to minimize the average distance between the DGZs and the center of the footprints which cover them. See §7.

Problem (6.2) leads to considerations of effective organization of the assigned DGZs to delivery systems. The collective behavior of the DGZs may be viewed as a game between automata. Pursuance of such a formulation would involve concepts discussed in [7].

7.0 A Method for DGZ Generation

Herein an alternative method for DGZ generation is suggested. Many of the techniques currently used will be common to this problem, such as the method of determining values extracted and varying the type of detonation, etc. The keystone of the suggestion is an algorithm which will iteratively generate aimpoints for collections of installations based on partitioning the installations into a distinct number of subsets in such a way that the aggregate "distance" from each installation to its associated aimpoint is minimal.

We assume that a set of complexes has been generated. Since there is no collateral damage between complexes, we shall generate aimpoints within a single complex; these will be independent of the aimpoints of any other complex. Assuming a flat earth in the complex, we can identify any installation i by its coordinates (x_i,y_i). We assume that the value extracted from installation i is known when the aimpoint, weapon yield and accuracy, type and height of burst, etc., are given. The value extracted also depends on the hardness of the installation.

7.1 Formulation and the Algorithm

For any fixed yield of weapon (tier), it is our objective to minimize the aggregate distance of each installation to its aimpoint. Let n be the number of DGZs and m be the number of installations. We obtain the problem

$$\text{minimize} \quad \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} \sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2} \qquad (7.1)$$

$$\text{subject to} \quad \sum_{j=1}^{n} \lambda_{ij} = 1 \qquad \forall i$$

$$\lambda_{ij} = 0 \text{ or } 1 \qquad \forall i,j$$

where (X_j, Y_j) are the coordinates of aimpoint j and λ_{ij} indicates whether installation i is covered by DGZ j (i.e., I(j)).

Cooper's algorithm [6] to solve problem (7.1) is

A) arbitrarily select n distinct aimpoints in the complex

B) assign each installation to its nearest aimpoint

C) designate the centroids of each of the n subsets created in B as the new aimpoints

D) iterate over step B and C until improvement is negligible.

An example of the application of this algorithm given by [2] is shown in figure 5.

In problem (7.1) we have assumed that n is fixed and known. However, n is unknown; so we must find the optimal choice of n. We suggest the following method:

A) Establish an initial value for n which should not exceed the lowest possible number of DGZs that could be expected to adequately cover the installations in the complex.

B) Apply Cooper's algorithm and determine the aimpoint locations.

C) Evaluate the total value extracted from the installations in the complex. The height of burst, etc., should be varied at this time to obtain the maximum possible value extracted.

D) Repeat steps B and C with an additional DGZ. Continue until the value extracted stabilizes.

A practical stopping should be based on consideration of the additional value extracted per added DGZ. Also, some acceptable value of damage, say 80% of the total value of the installations in the complex, can be applied as a stopping rule; thus, the optimal choice of n will be the minimum number of DGZs needed to obtain at least this acceptable value. In figure 6 the best choice of n is 3.

This proposed method of DGZ generation allows each complex and each tier to be considered independently. Completely nested layouts, layouts with nesting only in tier 1, or nonnested layouts may be constructed by this scheme. Complete nesting, however, is not recommended for the DGZ layout since it may not lead to an efficient extraction of value for some complexes.

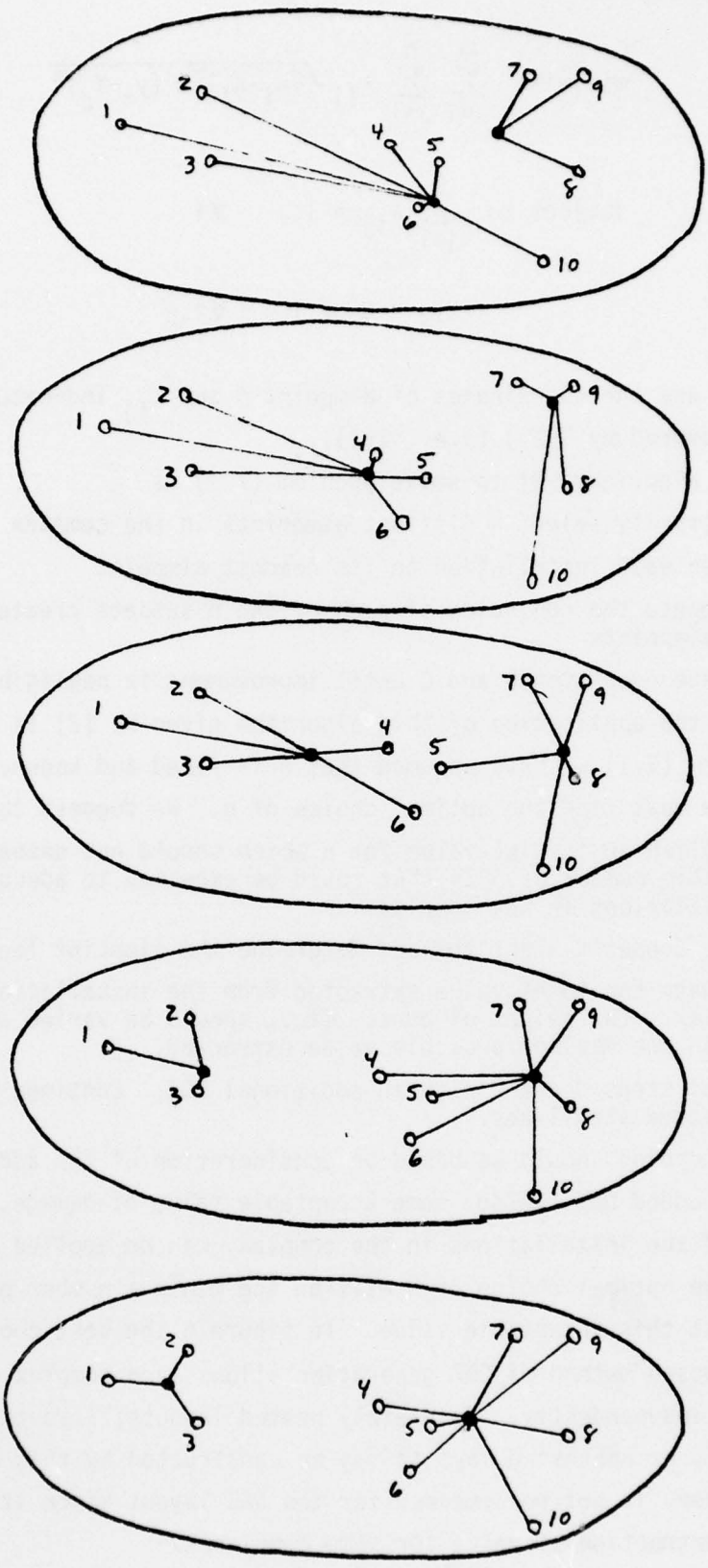


Fig. 5. Solution for m = 10, n = 2

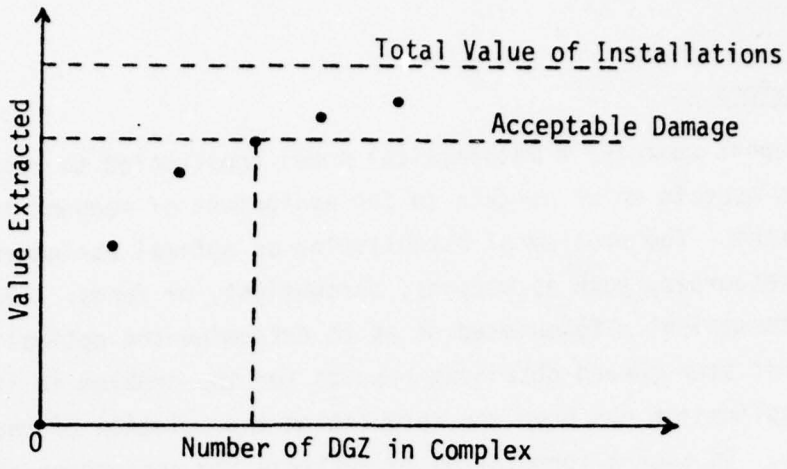


Fig. 6. Relationship of Value Extracted to Number of DGZs Used

7.2 Modfiications

Some possible weaknesses of the approach described in § 7.1 are its insensitivity to the hardness and value of installations when determining aimpoints. For example, suppose installation 9 in figure 5 has very high value, perhaps an order of magnitude higher than the next highest valued installation. The damage associated with installation 9 may be insufficient in terms of the above stopping rules. Of course, large values of n will correct this. However, we shall consider a method so that an aimpoint is attracted to installation 9. We divide installation 9 into k "dummy" installations, each with coordinates (x_9,y_9) and each with a point value $v_d = v_9/k$. The value of k is chosen to make v_d comparable to other installation values. An extreme use of this procedure would be to create "dummy" installations at each location, where v_d is based on the lowest valued installation in the complex. The creation of "dummy" installations allows greater weight to be placed on installations of very high value and insures that an aimpoint (centroid) will be placed close to its actual location. This follows from the form of the objective function of (7.1).

Another method of weighting certain installations is to modify the objective function of (7.1). For example, we might minimize

$$\sum_i \sum_j \lambda_{ij}\, h_i\, v_i \sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2}, \qquad (7.2)$$

where v_i is the value of installation i and h_i is a measure of the relative hardness of installation i.

8.0 Conclusions

This report contains a mathematical model constructed to represent some of the numerous aspects which pertain to the assignment of weapons in a tiered aimpoint system. The problem of establishing an optimal assignment arises due to limited resources, such as weapons, throwweight, or funds. Thus, the problem has been mathematically formulated so as to determine the optimal assignment.

The first step toward obtaining results for the problem in its most general practical application has been the formulation and solution of the pure assignment problem. In such a formulation we optimize the assignment of weapons given the supply and deployment of weapons and the layout and values extracted of the DGZs. A method of calculating the accessibility of each DGZ by each weapon is given. It is assumed that each weapon may be independently aimed to each DGZ in range. Due to the magnitude of any practical application, a computationally efficient method of solution is required. The associated linear programming problem (corresponding to the integer programming formulation) possesses a domain of feasible solutions which is rich in potential optimal solutions that are also integral. Thus, a branch-and-bound algorithm using a specially tailored (to take advantage of the particular structure) linear programming code as the evaluation and fathoming device is proposed as a computationally feasible method for finding optimal solutions to the pure assignment problem. In their own right, these solutions may be used as starting point for inclusion of new aspects and constraints for more general considerations of practical problems.

One such important additional constraint is the restriction of the weapons from a single ballistic missile delivery system to fall within its footprint. By simplification of the description of footprint and a transformation of the coordinate system, we have reduced the dispersion problem for MIRVs to its basic essentials. This canonical problem, however, remains unsolved.

Because of the complexity of the general problem, a sequence of more relevant models must be considered to converge to an applicable solution. Herein we have initiated a first stage. Increasingly detailed modeling requires a review of the results at every stage to judge progress and pertinence to the general problem. While intermediate solutions may not be feasible to the general problem, comparison with the presently used algorithms and their results is important to judge progress as well as to suggest further refinements.

Mathematical treatment of the target-rich and target-poor assignment problems has illuminated the modeling of several important related issues. Thus, the

techniques developed herein may lead to more effective allocation of weapons mix and deployment, especially when there are negotiable constraints (as in SALT) or cost considerations. Automation of the procedures developed would allow quick consideration of many of these issues and allow for sensitivity analysis of the results.

9.0 Recommendations

1. Work should continue on computer implementation of the proposed algorithm to solve the pure assignment problem and culminate in comparison against currently used methods.

2. Work should continue on the mathematical treatment of the MIRV dispersion problem.

3. Decisions on additional refinements should be made when 1 and 2 are completed.

4. Simulated data should be supplied so that results of the suggested DGZ generation algorithm may be compared to the present DGZ layout.

References

1. Cooper, L. (1967), "Solutions of generalized locational equilibrium models," Journal of Regional Science, 7, 1-18.

2. Garfinkel, R.S. and G.L. Nemhauser (1972), Integer Programming, Wiley, New York.

3. Hu, T.C. (1969), Integer Programming and Network Flows, Addison-Wesley.

4. Jewell, W.S. (1962), "Optimal flow through networks with gains," Operations Research, 10, 478-499.

5. Langley, R.W. (1973), "Continuous and integer generalized flow problems," doctoral dissertation, Georgia Institute of Technology.

6. Scott, A.J. (1971), Combinatorial Programming, Spatial Analysis and Planning, Methuen, London.

7. Tsetlin, M.L. (1973), Automaton Theory and Modeling of Biological Behavior, Academic Press, New York.

8. Wagner, H.M. (1969), Principles of Operations Research, Prentice-Hall, Englewood Cliffs, New Jersey.

9. Aeronautics 332, Course Text, USAF Academy, Colorado.

Acknowledgement

The author is grateful to the Air Force Systems Command and ASEE for support and administration of this summer research. The author is indebted to the Strategic Missile Division of Headquarters Air Force, Studies and Analysis, for supplying an interesting problem. The consultation and encouragement of Capt Bert Knight was extremely valuable.

The author is grateful to the Frank J. Seiler Research Laboratory for providing a cordial and stimulating work environment. It has been a pleasure and a professionally enriching experience to work among these very capable people. The author wishes to thank the personnel of the Applied Mathematics Branch for their assistance and encouragement during the course of the present work. In particular, thanks are due to Lt Col John Brush, Lt Roy Sikes, and Ms. Donna Weiss. The author wishes to express his sincere appreciation to Capt Salvatore J. Monaco for his valuable suggestions, consultation, and assistance.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM
sponsored by
THE AIR FORCE OFFICE SCIENTIFIC RESEARCH
conducted by
AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

ELECTROCHEMICAL AND LUMINESCENCE STUDIES

WITH

SPECIAL EMPHASIS ON MOLTEN SALT MEDIA

Prepared by:	Csaba P. Keszthelyi, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Chemistry Louisiana State University
Assignment:	
(Air Force Base)	USAF Academy
(Laboratory)	F. J. Seiler Research Laboratory
(Directorate)	Chemical Sciences
(Division(s))	Electrochemistry and Molecular Dynamics
USAF Research Colleagues:	Capt Charles L. Hussey Capt Larry P. Davis
Date:	August 12, 1977
Contract No.:	F44620-75-C-0031

ELECTROCHEMICAL AND LUMINESCENCE STUDIES WITH SPECIAL EMPHASIS ON MOLTEN SALT MEDIA

by
C. P. Keszthelyi

ABSTRACT

Molten salts have an important role in chemical technology, and developments in the past few years have dramatically reduced the temperature requirements. Whereas, for example, lithium chloride - potassium chloride melts have an operational temperature of several hundred degrees Celsius, some of the molten salts employed in this study have a liquidus temperature well below the freezing point of water.

The newness of these systems necessitates the acquisition of data delineating their basic properties, and the work described in this report concentrated on electrochemical and spectroscopic characteristics and applications. The alkylpyridinium halide/aluminum chloride melts have an electrochemical window that is primarily suited for cation studies, and over a dozen solute species covering a wide range of redox properties were investigated with satisfactory results. A unique property of some melts is the non-anodic oxidation of solutes having an oxidation potential less than +1.5 V vs SCE, followed by alkylation reactions via the pyridinium salt. The oxidizing strength of the melts was found to depend on the pyridinium salt: aluminum chloride ratio, opening the way for synthetic, preparative applications of these ambient temperature molten salt systems.

Spectroscopic work was aimed at both the characterization of the solute species that occur in the electrochemical context, and development of new laser systems primarily suited for guidance purposes. Part of this effort is represented by a graphical-output computer simulation program aimed at high frequency electrogeneration in light emitting electrochemical media. The optical window of the melts is strongly dependent on preparation and electrochemical preconditioning; it is generally well suited for visible transitions, but not for UV below 3500Å. A strong molecular association between solute and solvent shifted the fluorescence spectrum of several scintillator or laser dyes towards the red by as much as 500Å.

Non-bound ground state studies, involving dyes or inter-noble gas excimers/exciplexes, were at an exploratory stage at the end of the ten weeks. A remarkable extension of these molecular dynamics studies relates back to the solute aggregate states that were observed by ESR and UV-visible absorption techniques in these solvent electrolyte systems.

AD-A051 624 AUBURN UNIV ALA SCHOOL OF ENGINEERING F/G 5/2
1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM. VOLUME I.(U)
SEP 77 J F O'BRIEN F44620-75-C-0031
UNCLASSIFIED AFOSR-TR-78-0348 NL

6 OF 6
AD
A051624

END
DATE
FILMED
4-78
DDC

ACKNOWLEDGEMENT

The author is sincerely grateful to the Air Force Systems Command for support of this summer research. A deep appreciation is due to ASEE and Auburn University and in particular Mr. Fred O'Brien whose excellent administration of the project has made this summer more productive and enjoyable.

The author is indebted to the Electrochemistry and Molecular Dynamics Divisions of the Frank J. Seiler Research Laboratory for providing a cordial and helpful work environment. It has been a pleasure and a professionally gratifying experience to work with my immediate research colleagues, Capt Charles L. Hussey and Capt Larry P. Davis. Particular thanks are due to Colonel M. D. Bacon, Lt Colonel T. M. Tomaskovic, Lt Colonel B. A. Loving, Lt Colonel L. A. King, and Major J. K. Erbacher.

It has been enjoyable and profitable in working with the staff of FJSRL, namely Mr. F. C. Kibler, Jr. and Mr. J. L. Pflug. I wish to express my thanks to them.

LIST of FIGURES.

I. INTRODUCTION.

The essential phenomenon of electrochemistry is *charge transfer*, and this requires a conducting medium. Laws of transport for dilute solutions have been established for many years (1), still subject to modifications and improvement with the passage of time (2). Though the majority of electrochemical work, especially in academic institutions, is still carried out in dilute solutions both aqueous (3) and non-aqueous (4), many of the most important technological applications relate to concentrated solutions (5), molten salts (6), and solid electrolytes (7). While the elevated temperatures associated with the usual melts may be desirable under special circumstances such as presented by nuclear reactors (8), elevated temperature requirements more often present a hindrance from the experimental viewpoint, curtailing engineering and production (9).

Recent development in our laboratory of melts based on 1-alkylpyridinium chloride - aluminum chloride mixtures (10) has provided a technological as well as scientific breakthrough that has already led to important practical applications (11); apart from good stability and conductivity, these systems offer a unique temperature range *from boiling water to dry ice* slush. A principal advantage of these melts is their low RC value for double layer charging, known to be of critical importance in the design of an electrochemiluminescence (ECL) laser (12). It should be mentioned that there are a number of other important factors that play a role in ECL (13-30) which could not be adequately discussed within the present report, yet these may have a decisive impact upon a device containing one of our molten salts as solvent-electrolyte.

As of June 1977, virtually no data was available on spectroscopic properties of these melts. The extreme toxicity of some solvents like selenium oxychloride makes it important to continue the search for alternate solvents for laser applications, and some of these molten salts may provide a suitable medium. The unique properties of these solvents, on the other hand, may require a reexamination of current theories dealing with lasers (31-37).

II. OBJECTIVES.

As evident from the Introduction which surveyed the state of knowledge on ambient temperature 1-alkylpyridinium chloride - aluminum chloride melts, there is a great lack of basic information due to the newness (10,11) of these systems. Consequently one of the objectives of the present work was to do exploratory studies on various chemical and physical properties of the melts. The solutes were chosen from among previously investigated molecules, so that the observation would allow ready comparison of the molten salts with commonly used electrochemical and spectroscopic solvents. Within the broader context of the framework of several years hence, the accumulated information should be of value in the development of an electrochemically pumped laser, i.e. one in which a flashlamp excitation source is replaced by a small DC battery. Such a device would be useful in communication and guidance purposes, rather than high power applications. One of the planned summer activities in this area was computer modelling, including graphical representations, of relevant homogeneous as well as heterogeneous charge transfer processes at and in the vicinity of the electrodes, utilizing the USAF Academy computer facilities. A considerable amount of relevant information is already in the research literature (38-41).

Thermal batteries have a number of important defense applications, and part of the 10 week research program was to be devoted to this topic. In view of the formidable background already assembled, the activity was to be less exploratory, dealing mainly with improvement of performance and reliability of chromium based thermal batteries, and replacement of the critical item chromium through substition with comparable materials.

Laser dyes are subject to decomposition upon prolonged use, and this would be a drawback of an ECL pumped laser made along the line of previous attempts (42). Use of optically active molecules that do not decompose in the usual chemical sense would seem especially important for high power applications. A general lack of information of the *molecular dynamics* type appears to be characteristic of several important topics related to the above, particularly concerning non-bound ground state dimers of both laser dyes and *noble gas excimers*, *exciplexes*, and *hypermolecules*. Some exploratory work in this area was also among the planned summer itinerary.

III. EXPERIMENTAL

The outstanding instrumental facilities available at the Frank J. Seiler Research Laboratory will no doubt be an incentive for future participants in the Program also. Figure 1 summarizes in a block-diagram schematic the principal instrumental set-up used for the combined electrochemical/spectroscopic study, utilizing the following equipment: Keithley 244 (high voltage power supply), Oriel 8540 (arc lamp power supply), Oriel 6140 (1000 watt Hg-Xe arc lamp), two McPherson monochromators (0.5 m, 1 m), PAR-192 (Princeton Applied Research Corporation, variable frequency light chopper), PAR-181 (current sensitive pre-amplifier, 6 decades range), HI-2000 (Houston Instruments "Omnigraph" XY-timebase recorder), Jarrel-Ash 82-410 (1/4 m monochromator), Tektronix-556 (dual beam oscilloscope), PAR-173 (potentiostat, equipped w. PAR-179 digital coulometer), PAR-175 (universal electrochemical programmer). For optical alignment a downgraded performance He-Ne laser (Model 132 by Spectra-Physics) was used.

UV-visible absorption studies were carried out with a Cary Model 15 Spectrophotometer equipped with a Cary Model 1115 Repetitive Scan Accessory. Difficulties involving the collection optics of the arrangement of Figure 1 made it desirable to perform some of the low temperature fluorescence studies with a Perkin-Elmer Model 204 Fluorescence Spectrophotometer. The superior resolution of the Jarrel-Ash components was not indispensable for the condensed phase fluorescence observations, and reproducibility of the rigid geometry commercial equipment was a positive asset.

The controlled temperature cell holder, equipped with slits for 90° fluorescence, is shown in Figure 2. A Vulcan 100 watt cartridge heater seated at the lower part of the aluminum block was controlled by a variable voltage transformer.

The solutes that were investigated included 9,10-diphenylanthracene (DPA), Rhodamine-B, 5,6,11,12-tetraphenylnaphthacene (rubrene), 2,5-diphenyloxazole (PPO), 2,5-diphenyl-1,3,4-oxadiazole (PPD), thianthrene, europium oxide, 9,10-dimethylanthracene (DMA), N,N,N^1,N^1-tetramethyl-p-phenylenediamine (WB), and 1,4-bis-(5-phenyloxazol-2-yl)-benzene (POPOP), in molten ethylpyridinium bromide (EtPyBr), 1-methylpyridinium chloride/aluminum chloride (MePyCl/$AlCl_3$, 1-2 m/o), and 1-butylpyridinium chloride/aluminum chloride (BuPyCl/$AlCl_3$). Conventional solvents employed in the study were spectroscopic or reagent grade, and included benzene, toluene, hexane, acetonitrile (ACN), N,N-dimethylformamide (DMF), tetrahydrofuran (THF), and dichloromethane ($MeCl_2$).

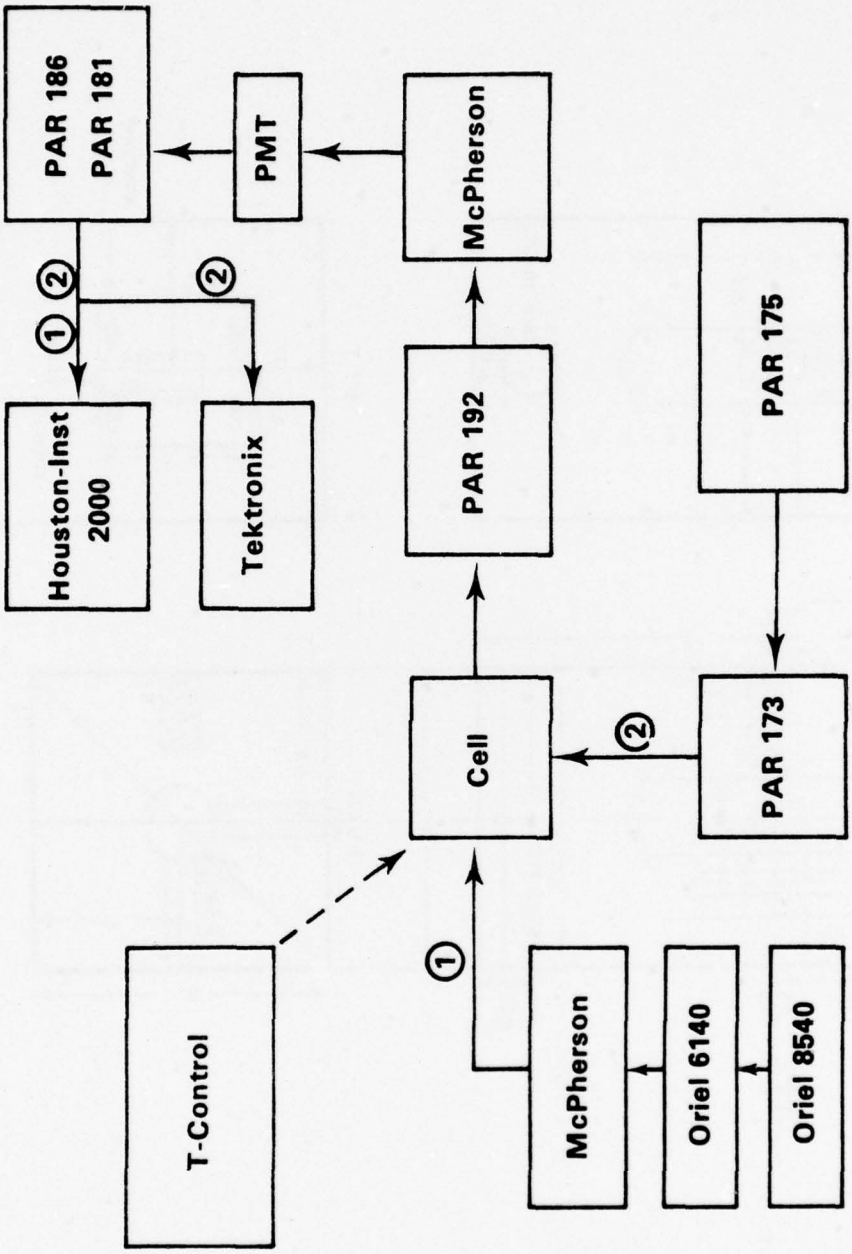

PAR 186
PAR 181
PMT
McPherson
PAR 175
Houston-Inst
2000
Tektronix
PAR 192
Cell
PAR 173
T-Control
McPherson
Oriel 6140
Oriel 8540
1
2

Figure 1

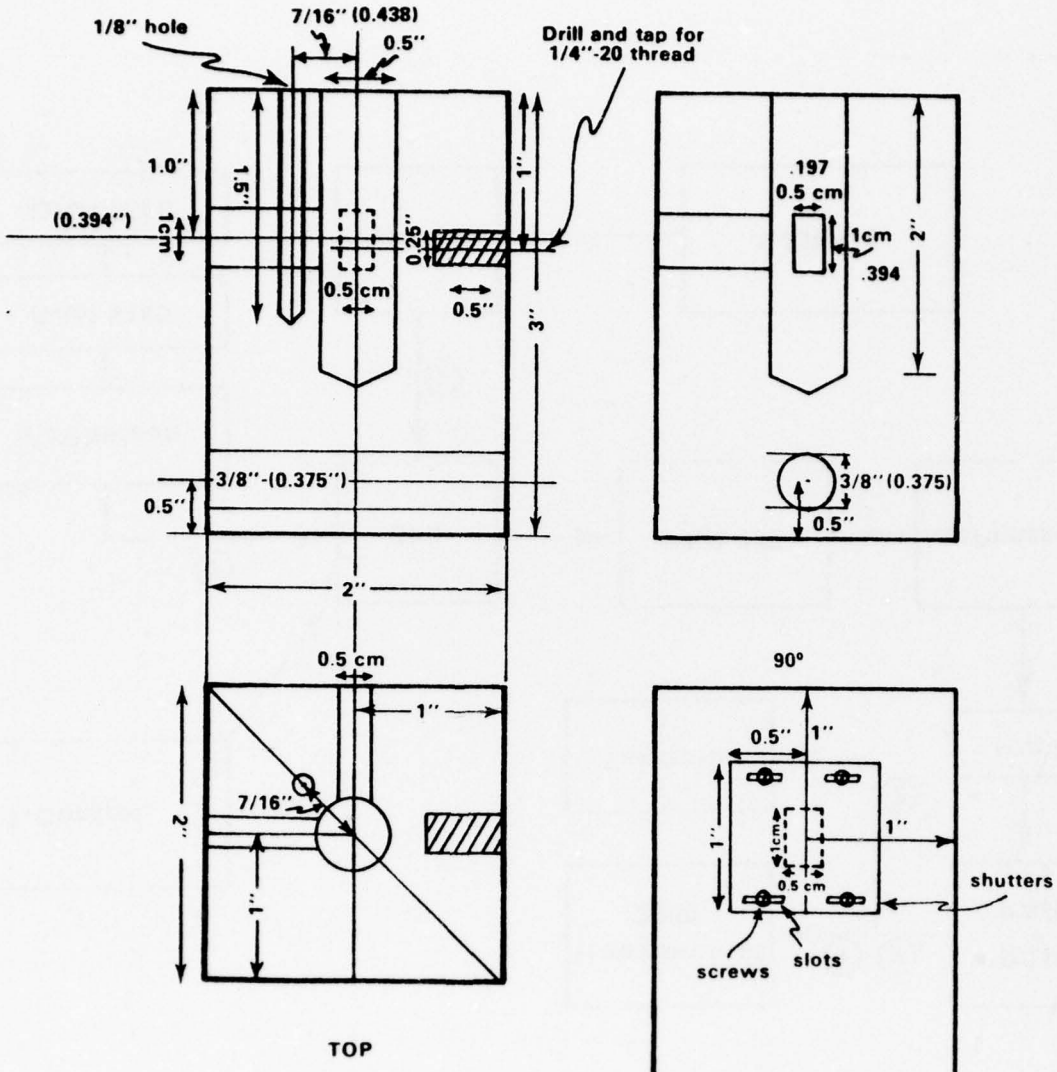

1/8" hole
7/16" (0.438)
0.5"
Drill and tap for
1/4"-20 thread
1.0"
(0.394")
1cm
1.5"
0.25"
1"
0.5 cm
0.5"
3"
3/8"-(0.375")
0.5"
2"
0.5 cm
1"
7/16"
2"
1"
TOP
.197
0.5 cm
1cm
.394
2"
3/8"(0.375)
0.5"
90°
0.5"
1"
1"
1cm
1"
0.5 cm
shutters
screws
slots
SHUTTER ARRANGEMENTS

Figure 2

IV. RESULTS AND DISCUSSION

Under the necessary precautions to exclude air and moisture from the electrochemical cell, a cyclic voltammogram (potential excursion: +1.5V to +2.1V to +0.9V to +1.5V (vs. Al)) of thianthrene in 1-butylpyridinium chloride/aluminum chloride (1:2 m/o) at a wolfram wire electrode shows (Figure 3) stable TH^+ and TH^{++} formed in reversible Nernstian charge transfer steps (ip_a/ip_c was unity for scan rates of 0.010V/sec to 0.500V/sec; $Ep-E_{1/2}$ and $Ep_{/2}-E_{1/2}$, 30 mV), akin to Mamantov's (43) high temperature melt results. Figure 4, also a cyclic voltammogram of TH in $BuPyCl/AlCl_3$, shows some novel chemical effects; when the potential excursion is from +1.5V to +0.9V to +1.5V (vs. Al), it is revealed that TH^+ is already present in the solution, hence a non-anodic oxidation of TH has taken place. This can be interpreted under our rigorous experimental conditions as a reaction between TH and the alkylpyridinium cation, shown by the first *insert* on page 20. (The second *insert* is a tentative disproportionation scheme.)

A subsequent disproportionation of the alkylpyridine radical to yield pyridine, and octane, or butylated products, still needs further verification. The observed non-anodic oxidation of WB, nubrene, DPA and DMA suggests that the observations are of a broad significance -- solutes whose $E^\circ_{R/R+}$ is below +1.5V vs SCE undergo the redox reaction in these melts, whereas the other solutes we have investigated, e.g. PPD ($E^\circ_{R/R+} \simeq 2.2V$ vs SCE), are stable for many days in these solute-electrolytes.

The thianthrene cation has been extensively studied by Shine et al. (44), and we have observed an ESR signal as expected (43-45) at -120°C, shown in Figure 5. Due to the very high dielectric constant of the liquid melts, liquid phase measurements were not feasible using routine techniques.

The characteristic purple-violet color of TH^+ develops within minutes after TH is placed in the melts, and absorption spectra revealed the TH^+ absorption peak at 534 nm. The optical window of these melts is not suited for directly observing the TH absorption disappearance ($\varepsilon \simeq 0.05$ liters/mole-cm from 3500Å to 7000Å). Temperature dependence of the TH^+ absorption spectrum is in agreement with the results of

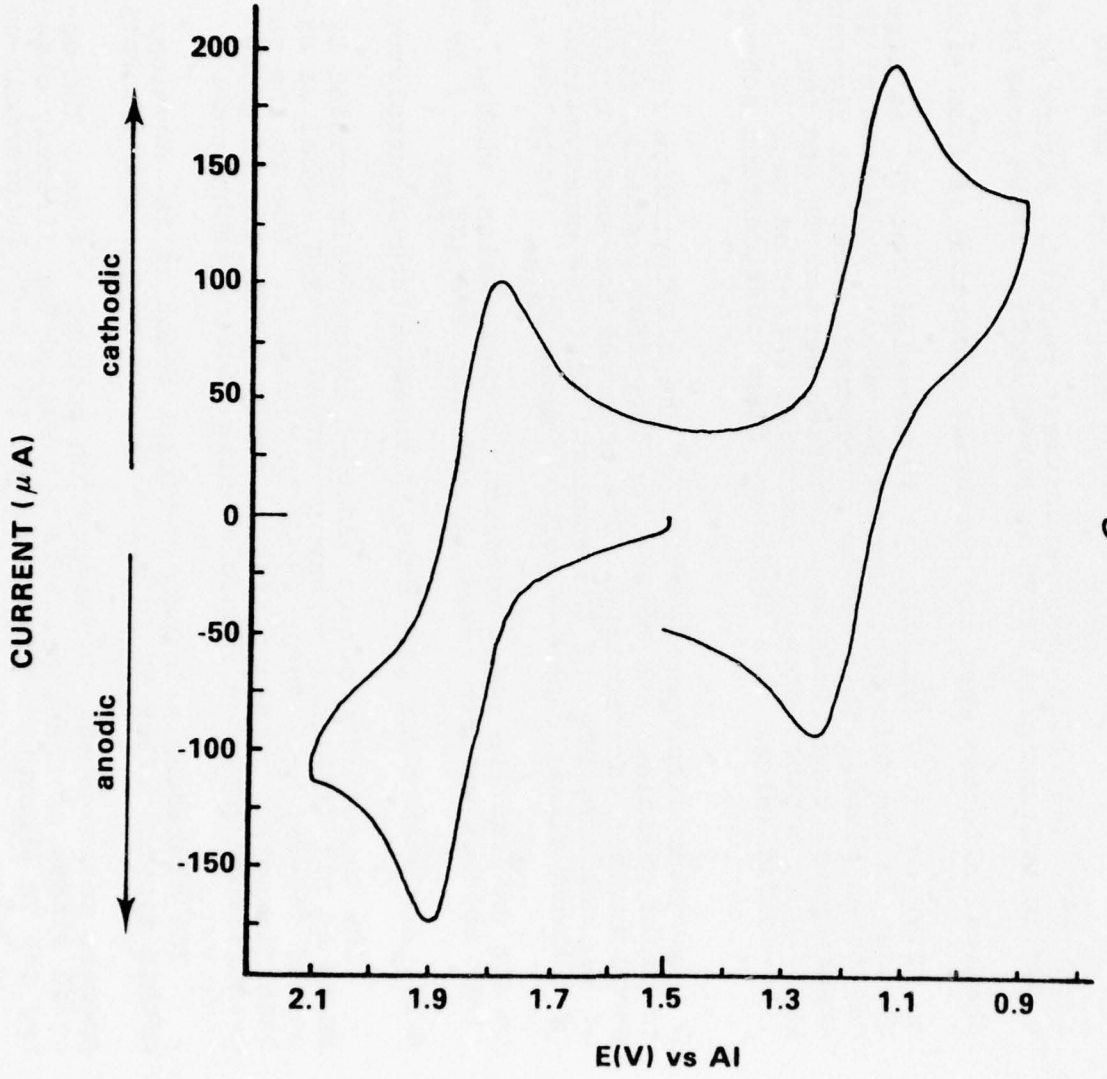
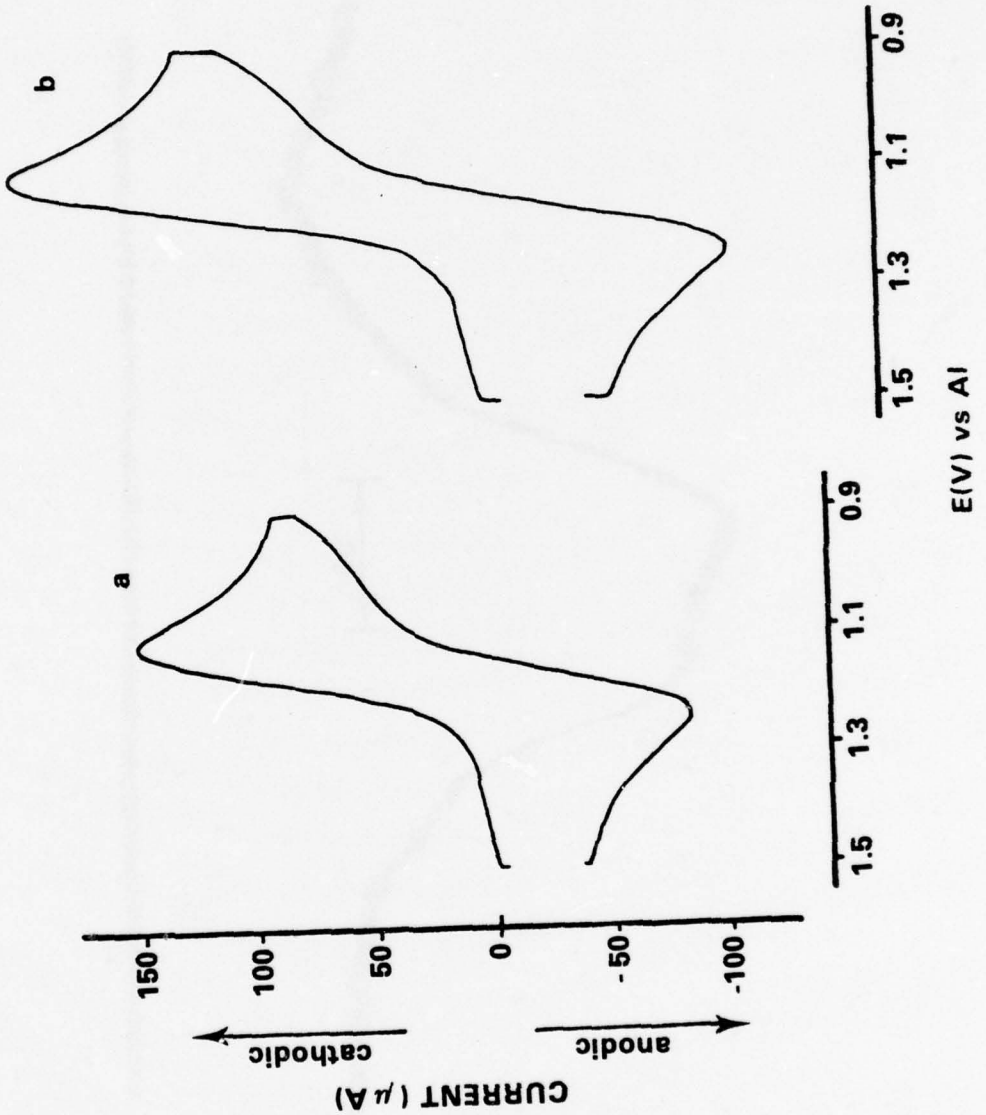

CURRENT (μA)
cathodic
anodic
200
150
100
50
0
-50
-100
-150
2.1
1.9
1.7
1.5
1.3
1.1
0.9
E(V) vs Al

Figure 3

Figure 4


b
a
150
100
50
0
-50
-100
cathodic
anodic
CURRENT (μA)
1.5
1.3
1.1
0.9
E(V) vs Al

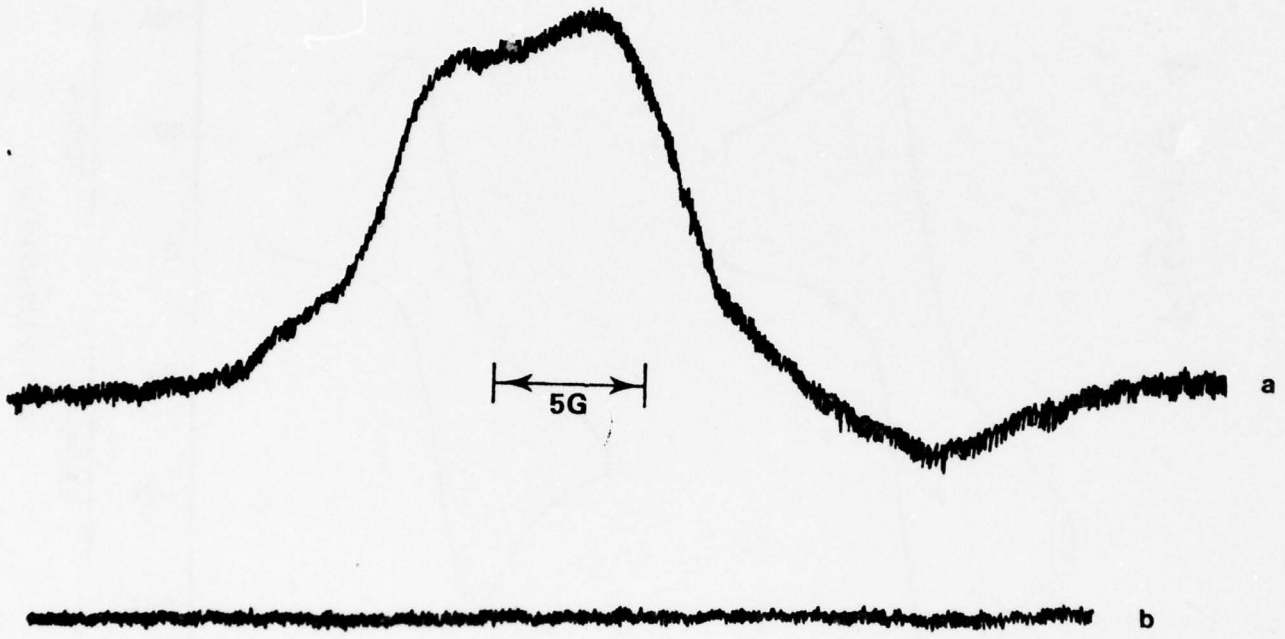


Figure 5

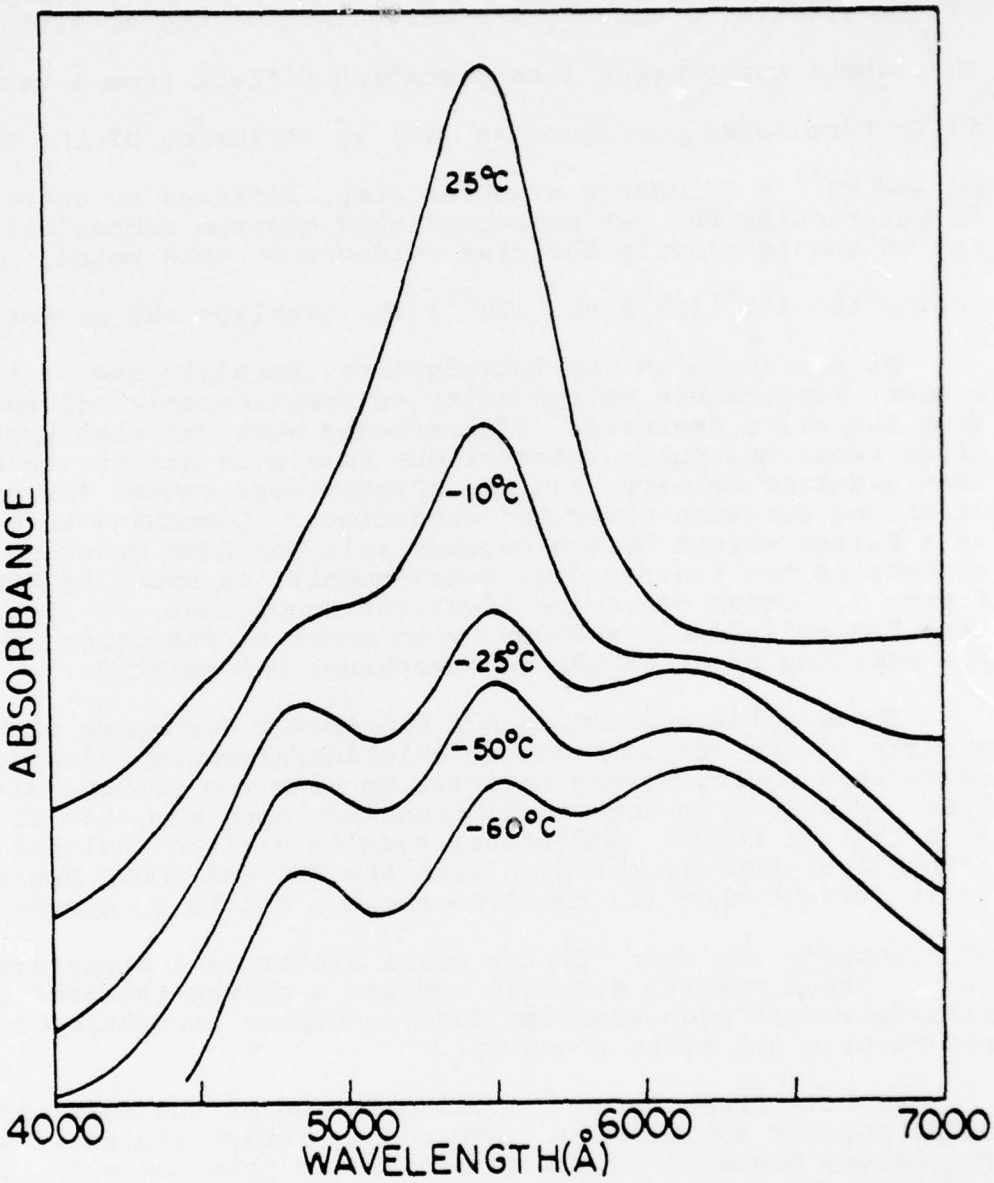

25°C
-10°C
-25°C
-50°C
-60°C
ABSORBANCE
4000
5000
6000
7000
WAVELENGTH(Å)

Figure 6

Szwarz et al. (45), and is shown in Figure 6. Interpretation of the temperature dependence, in view of the reversibility of the effects, could not logically involve decomposition of TH^+; where the present interpretation differs from a carefully formulated previous one (45) is inclusion of the $TH^+ + TH^+ \rightleftharpoons TH^{++} + TH$ charge transfer step, *followed* by aggregation. An interesting but yet unaccomplished quantum mechanical calculation should provide decisive evidence on this point, by comparing the $(TH^+)(TH^+)$ vs. $(TH^{++})(TH)$ overlaps and geometries.

As mentioned in the Introduction, possible use of the ambient temperature molten salts as spectroscopic solvents has some lucrative features. Fluorescence work can also provide clues about molecular interactions in a more direct manner than electrochemistry, and the solutes were tested for excitation and emission under 90° conditions. Quenching as well as a filter effect by the organic salt can have drastic effects on the fluorescence measurements, as shown by DPA in Figure 7. Under otherwise identical conditions, it is seen that DPA emission is reduced by an order of magnitude upon the addition of 5% EtPyBr to a methanol-DPA solution.

Those solutes which do not completely decompose within minutes in the alkylpyridinium chloride/aluminum chloride melts show a very strong interaction with the solvent-electrolyte. Figure 8 is the PPO fluorescence spectrum, and Figure 9 is that of POPOP. Additional results will be included in a Technical Report, but even from the PPO and POPOP behavior it is evident that (1) the fluorescence shifts by approximately 500Å towards the red; (2) the usual vibrational structure is lost. These results strongly suggest a charge-transfer complex formation even when the donor-acceptor energetics do not favor a net redox reaction.

We have also tested the possible use of the melts in a mixed solvent application. DMF reacts rather violently with the melts, hence it should be avoided in this context. On the other hand THF, ACN, $MeCl_2$ and other halogenated paraffins gave only a mild heating effect upon mixing. It would have been of interest to quantitatively evaluate the heats of mixing, and we have plans of a Joule-calorimetric study as a follow up on this summer work.

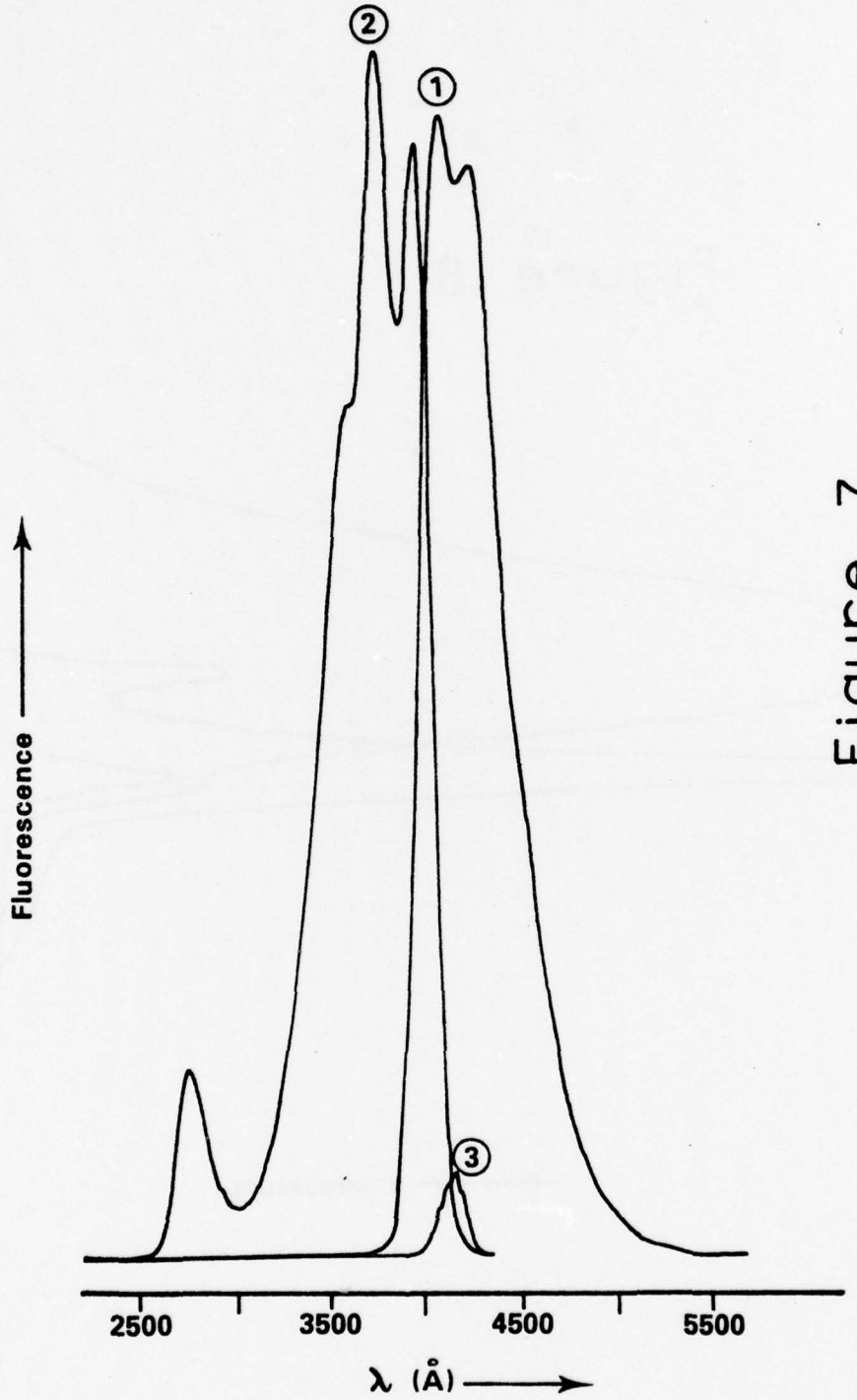


Figure 7

Figure 8

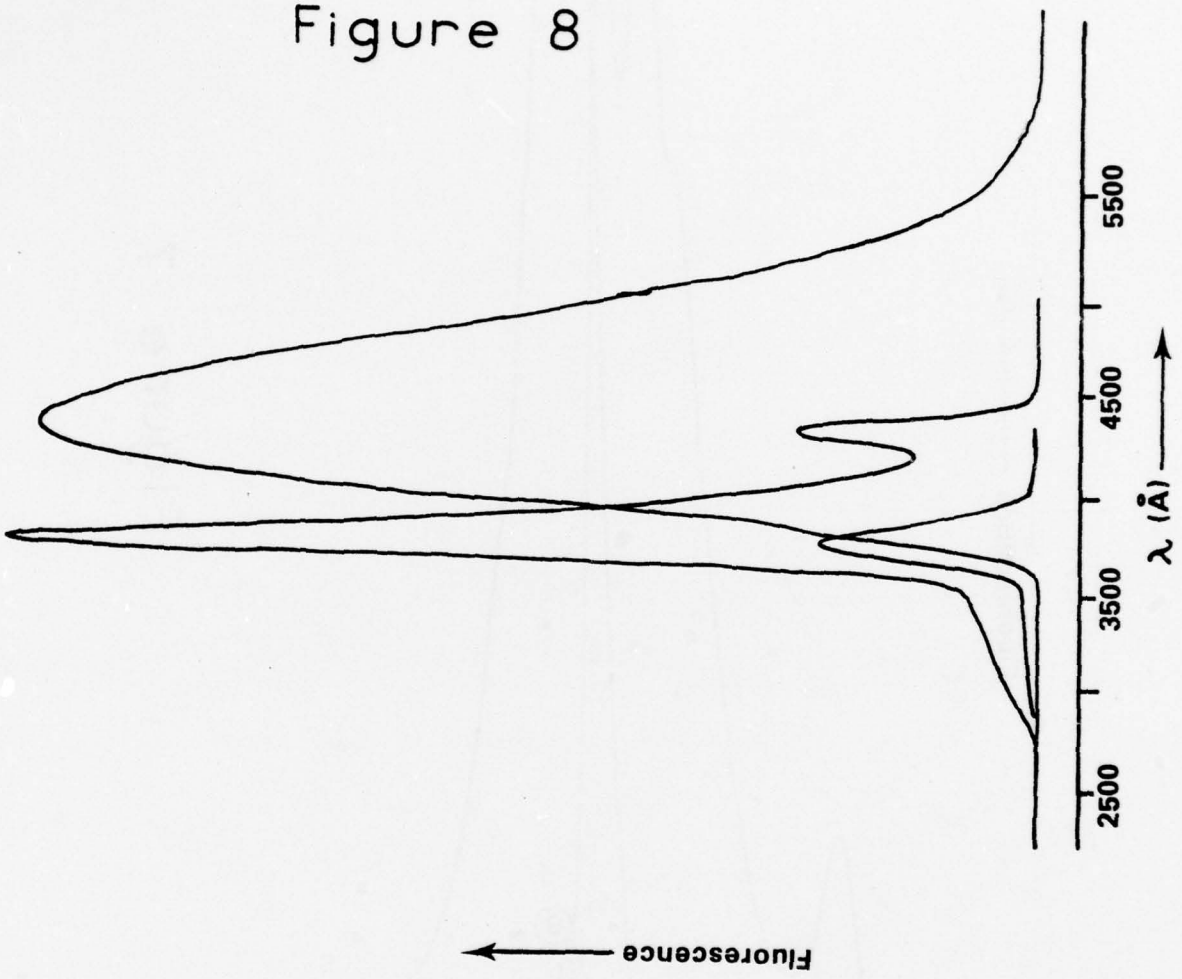
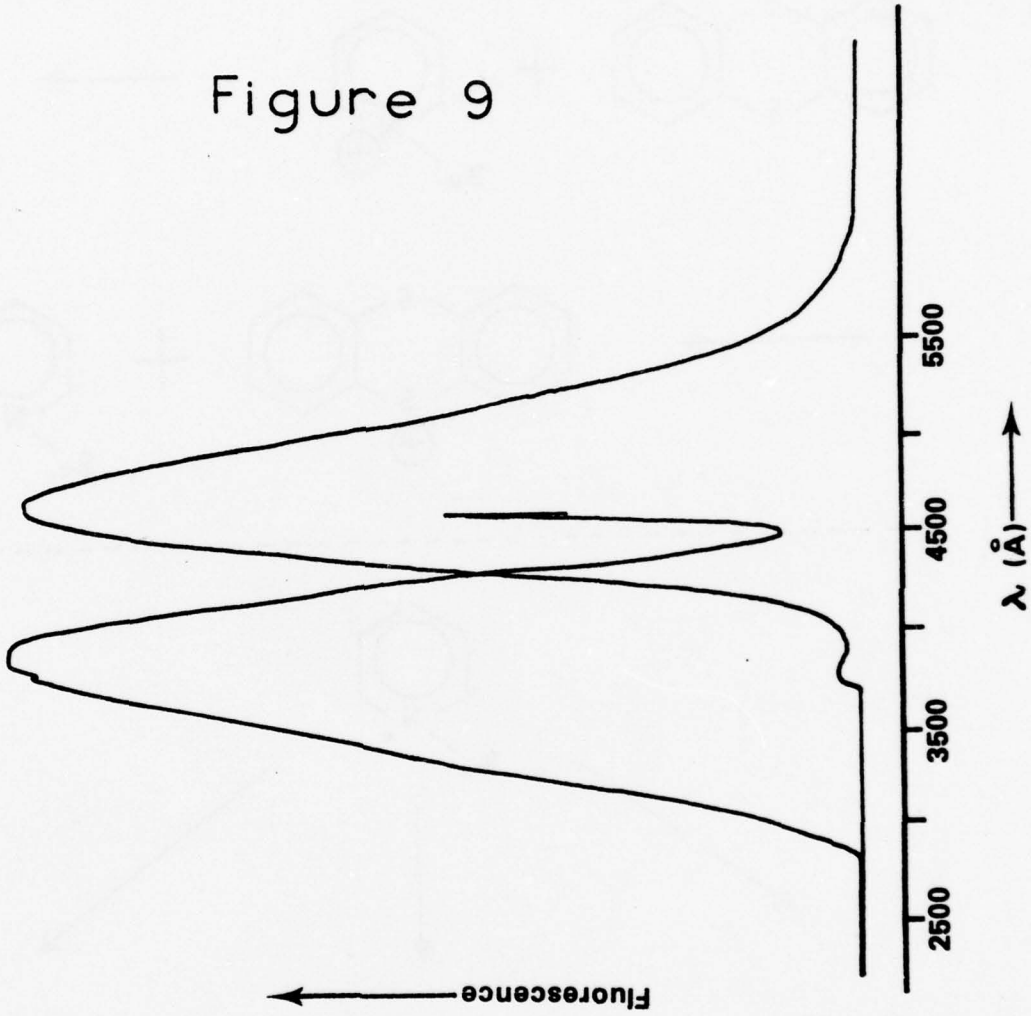
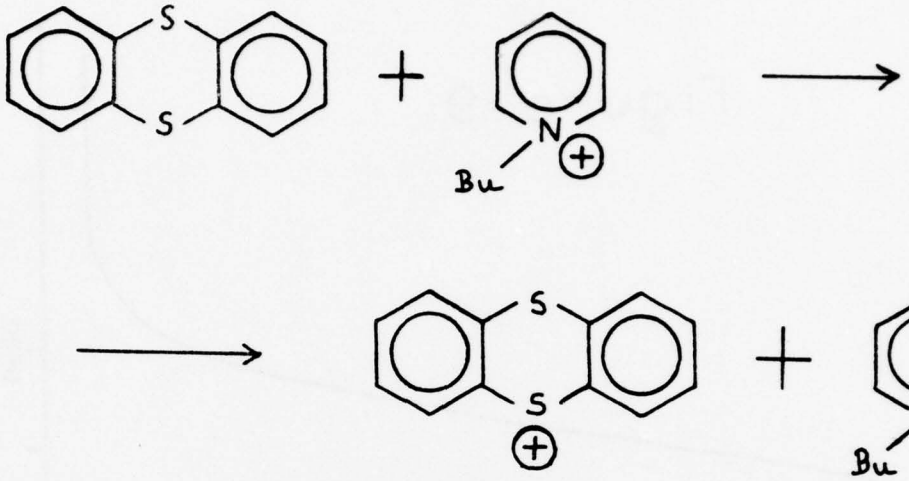
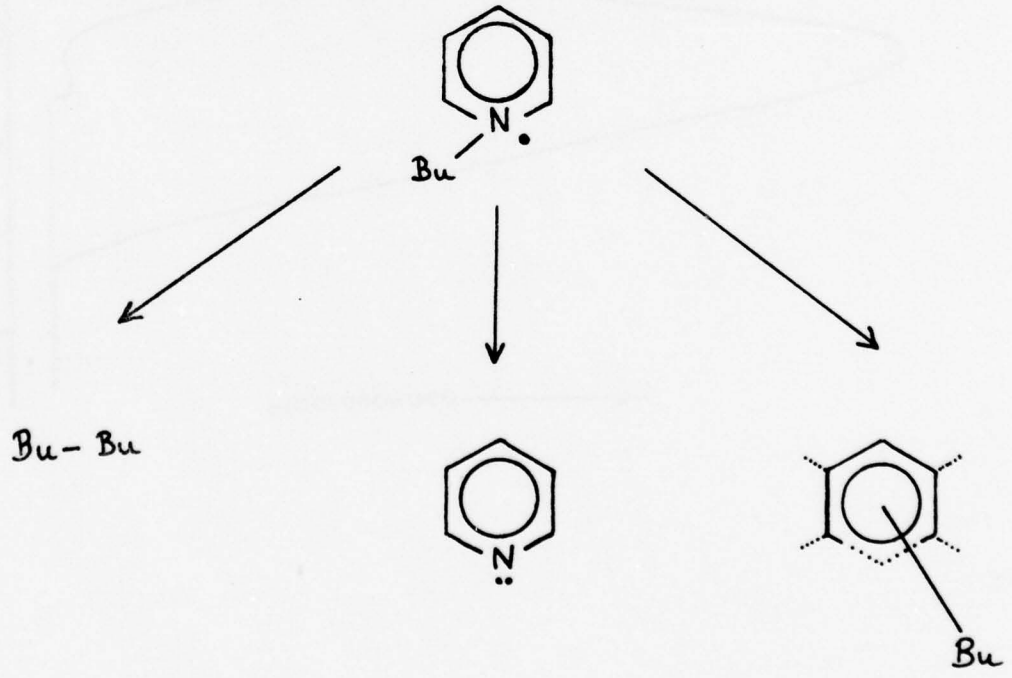


Figure 9



Computerized graphical representations pertaining to electrode kinetics, double-layer changing, and concentration profiles will be included in the Technical Report mentioned above.





V. CONCLUSIONS AND RECOMMENDATIONS

The present study has accomplished some exploratory work on ambient temperature molten salts based on 1-alkylpyridinium chloride - aluminum chloride mixtures. Developed at FJSRL prior to the commencement of this summer, these systems present a very significant scientific and technological breakthrough, and may well make a major contribution to the nation's energy needs.

The melts provide a generally suitable medium for anodic oxidations. They were also found to have a non-anodic oxidizing capability, similar to that of sulfuric acid for thianthrene. This suggests their possible extended use for cation spectroscopic studies. A further interesting detail uncovered in the 10 week summer activity was that the oxidizing strength of the melts depends upon the organic salt: aluminum chloride ratio. An explanation is provided by considering the 'bare' charge of the pyridinium cation due to the $Al_2Cl_6^- + Cl^- \rightleftharpoons Al_2Cl_7^-$ equilibrium which lies far to the right. Generation of H^+ in the subsequent redox step is explained by alkylation of the solute via the alkylpyridinium cation.

Optical properties of the melts were also investigated in the summer work, with favorable conclusion for the visible, but not for the UV, range of the spectrum. Scattering appears to be a major problem at high photon density.

Further work on the characterization of both the electrochemical and spectroscopic properties should be carried on. Several types of electrode materials should be tested with several solutes mentioned in the Experimental Section. Due to their convenient temperature range, the melts may provide a means for improved secondary batteries. Spectroscopic work in the visible and ESR region, including temperature dependence, promises further interesting results. Dimeric species of cation aggregates, compared with neutral solute dimers, present an important area of investigation for molecular dynamics, related both to electrochemistry and lasers. Non-bound ground state studies, perhaps using *ab initio* methods incorporating hyperpolarization, need to be extended from dye molecules to noble gases and their compounds.

REFERENCES

(1). M. Planck, Ann. Physik Chemie, NF, 39, 161 (1890).

(2). C.P. Keszthelyi, Proc. Louisiana Acad. Sci., 38, 107 (1975).

(3). L.M. Southwick, G.H. Willis, P.K. Dasgupta, and C.P. Keszthelyi, Anal. Chim. Acta., 82, 29 (1976).

(4). C.P. Keszthelyi, J. Amer. Chem. Soc., 96, 1243 (1974).

(5). S. Hills, J. Electrochem. Soc., 108, 810 (1961).

(6). H.A. Laitinen and W.V. Ferguson, Anal. Chem., 29, 4 (1957).

(7). C. Wagner, J. Chem. Phys., 21, 1819 (1953).

(8). B.F. Hitch and C.F. Baes, Jr., Reactor Chemistry Div. Annual Progr. Rept., Dec. 31, 1966, USAEC Rept. ORNL-4076 (1967).

(9). C.P. Keszthelyi and A.J. Bard, J. Electrochem. Soc., 120, 241 (1973).

(10). U.S. Patent (pending).

(11). C.L. Hussey, J.C. Nardi, L.A. King, and J.K. Erbacher, J. Electrochem. Soc., 124, 1451 (1977).

(12). C.P. Keszthelyi, Appl. Optics, 14, 1710 (1975).

(13). A.E. Chandross, Trans. N.Y. Acad. Sci. (Ser. 2), 31, 571 (1969).

(14). C.P. Keszthelyi, H. Tachikawa, and A.J. Bard, J. Amer. Chem. Soc., 94, 1522 (1972).

(15). N.E. Tokel, C.P. Keszthelyi, and A.J. Bard, J. Amer. Chem. Soc., 94, 4872 (1972).

(16). C.P. Keszthelyi, J. Electrochem. Soc., 120, 39C (1973).

(17). C.P. Keszthelyi, N.E. Tokel-Takvoryan, H. Tachikawa, and A.J. Bard, Chem. Phys. Ltrs., 23, 219 (1973).

(18). C.P. Keszthelyi and A.J. Bard, Chem. Phys. Ltrs., 24, 300 (1974).

(19). C.P. Keszthelyi and A.J. Bard, Extended Abstr., 72, 275 (1972).

(20). C.P. Keszthelyi and A.J. Bard, J. Org. Chem., 39, 2936.

(21). C.P. Keszthelyi, Acta Phys. et Chem. Szeged., 20, (4), 365 (1974).

(22). C.P. Keszthelyi, N.E. Tokel-Takvoryan, and A.J. Bard, Analytical Chemistry, 47, 249 (1975).

(23). C.P. Keszthelyi, Bull. Chem. Soc. Japan, 48, 1083 (1975).

(24). C.P. Keszthelyi, Spec. Ltrs., 8, 25 (1975).

(25). C.P. Keszthelyi and A.T. Kessel, Proc. Louisiana Acad. Sci., 38, 98 (1975).

(26). C.P. Keszthelyi, J. Electrochem. Soc., 122, 1643 (1975).

(27). C.P. Keszthelyi, J. Electrochem. Soc., 122, 1642 (1975).

(28). "On the Quantum Efficiency of Electrogenerated Chemiluminescence", A.J. Bard, C.P. Keszthelyi, H. Tachikawa, and N.E. Tokel, in "CHEMILUMINESCENCE and BIOLUMINESCENCE", M.J. Cormier, D.M. Hercules, and J. Lee, eds., Plenum Press, London - New York, 1973, pp. 197-205.

(29). C.P. Keszthelyi, Croatica Chem. Acta, 48, 25 (1976).

(30). C.P. Keszthelyi, A.T. Kessel, W.B. Edwards, and G.D. Clark, Proc. Louisiana Acad. Sci., 39, 65 (1976).

(31). C.P. Keszthelyi, Spec. Ltrs., 7(1), 19 (1974).

(32). C.P. Keszthelyi, Spec. Ltrs., 7(1), 27 (1974).

(33). C.P. Keszthelyi, Acta. Phys. et Chem. Szeged, 19(3), 221 (1973).

(34). C.P. Keszthelyi, Spec. Ltrs., 7(2), 85 (1974).

(35). C.P. Keszthelyi, Acta Phys. et Chem. Szeged, 20, (4), 365 (1974).

(36). C.P. Keszthelyi, Spec. Ltrs., 8, 341 (1975).

(37). C.P. Keszthelyi, Spec. Ltrs., 8, 931 (1975).

(38). W.V. Childs, J.T. Maloy, C.P. Keszthelyi, and A.J. Bard, J. Electrochem. Soc., 118, 874 (1971).

(39.) C.P. Keszthelyi, Spec. Ltrs., 7(8), 409 (1974).

(40.) C.P. Keszthelyi and W.B. Edwards, Proc. Louisiana Acad. Sci., 38, 102 (1975).

(41.) C.P. Keszthelyi, Proc. Louisiana Acad. Sci., 38, 14 (1975).

(42.) J.T. Bowman, Ph.D. Thesis, The University of Texas at Austin, 1970.

(43.) K.W. Fung, J.W. Chambers, and G. Mamantov, J. Electroanal Chem., 47, 81 (1973).

(44.) H.J. Shine, C.F. Dais, and R.J. Small, J. Org. Chem., 29, 21 (1964).

(45.) M. de Sorgo, B. Wasserman, and M. Szwarz, J. Phys. Chem., 76, 3468 (1972).

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

AUBURN UNIVERSITY AND THE OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

TESTS OF TWO HYPOTHESES CONCERNING THE OPTICAL PROPERTIES OF CRYODEPOSITED GASES

Prepared by:	William Barlow Newbolt, Ph.D
Academic Rank:	Professor of Physics
Department and University:	Department of Physics Washington & Lee University Lexington, VA 24450
Assignment:	
Air Force Base:	Arnold Air Force Station, Tenn
Laboratory:	Arnold Engineering Development Center
Directorate:	Test Engineering
Division:	Research
USAF Research Colleague:	Jeffrey A. Roux, Ph.D (ARO, Inc, AEDC operating contractor)
Date:	August 5, 1977
Contract No.:	F44620-75-C-0031

TESTS OF TWO HYPOTHESES CONCERNING THE OPTICAL PROPERTIES OF CRYODEPOSITED GASES

by

William Barlow Newbolt

ABSTRACT

The measurements which have been made at AEDC of the optical transmission of cryodeposited gases and gas mixtures on a cryogenically cooled germanium substrate have permitted checking two hypotheses concerning these deposits. The first hypothesis tested was that the mole fraction of an infrared-active molecule in a mixture of gases is equal to the ratio of its band strength in the cryodeposited mixture to its band strength in the pure cryodeposit. The hypothesis appears to hold for mixtures of one infrared-active molecule with nitrogen which is not infrared active. It does not hold for mixtures which contain more than one infrared-active molecule.

The second hypothesis checked was that the real part of the index of refraction can be calculated from the imaginary part with Subtractive Kramers-Kronig.(1) This hypothesis has been verified with great accuracy away from absorption bands. In the absorption bands, particularly the very sharp band of CO, the fluctuations in the index values from the Kramers-Kronig formalism are not as strong as the fluctuations in the values from the least-squares program.

Acknowledgements

The author is grateful to the people of this area and AEDC for their friendliness and hospitality. Dr. Lynn Laster and Mr. Fred O'Brien are to be congratulated for their attention to the administrative details which has made the program run smoothly. The helpfulness of Dr. Herman Scott in all phases of this work is gratefully acknowledged. His knowledge of optics in general and infrared spectroscopy in particular was very valuable to the author.

Thanks are due to another group, the employees of ARO, Inc, who have been consistently helpful. In particular, I would like to thank Dr. Jeffrey Roux and Mr. Bob Wood who are building a data base on the optical properties of condensed gas which supported this study and will support others in the future. They helped the author learn how the Fourier transform spectrometer works and used the least-squares program to get the data needed for band integration. I am grateful for Dr. John Pipes', as yet unpublished, technical report[5] which gave the author an understanding of the analytical model which was used to interpret cryodeposit transmission data. Dr. Allie Smith is due my thanks for allowing the author to share his office. He managed to watch without alarm as the computer printouts and card boxes gradually inundated his office.

Before I leave the subject of computing, I must thank Mr. Charlie Lowe who helped the author "clean up his formats," and the efficient key punch operators in the data entry office.

The author wishes to thank Mrs. Kathy White who typed this paper.

List of Appendices

Objective

The objective of this study has been to use the newly acquired data base on the optical properties of cryodeposited gases to develop a method of analysis which could be used to determine the mole fraction of infrared-active molecules in a cryodeposit mixture. The hypothesis used was that the strength of an absorption band was a linear function of the concentration of the molecule causing the absorption.

Since the Subtractive Kramers-Kronig[4] formalism is frequently used to calculate the index of refraction of a sample from absorption measurements, and since the least-squares program of Roux and Palmer[6] gives an independent value of the index of refraction, an objective of this work has been to compare the least-squares results with those from the Kramers-Kronig program.

Test of the First Hypothesis

In his final report on his summer research activity, Dr. Kent Palmer suggested that it might be possible to use the infrared absorption of a cryodeposit as a method of analysis for its constituents. Although he outlined several methods of analysis, an inspection of the data leads one to suspect that the most promising approach is to assume that the band strength of a molecule in a mixture is a linear function of the concentration of the infrared-active molecule in the mixture. Specifically,

$$\text{Mole Fraction} = \frac{\int_{\text{Band}_i} k_{mix}(\omega)\, d\omega}{\int_{\text{Band}_i} k_{pure}(\omega)\, d\omega}$$

Where k(w) is the imaginary part (absorption index) of the complex index of refraction of the cryodeposit.[3]

When the author came down to AEDC right after Memorial Day, he began to work with Bob Wood and Jeff Roux in an effort to learn how the transmission data are taken with the Fourier transform spectrometer. At that time much of the experimental data studied here had been taken. The least-squares program written by Roux and Palmer[6] had been used to find the values of the real and imaginary parts of the index of refraction at points 10 cm^{-1} apart over the range from 700 cm^{-1} to 3700 cm^{-1}. However, there was data in the data tapes spaced every 2 cm^{-1} and it was evident that these data would have to be used before any accurate band integrals could be calculated. Therefore, Jeff Roux began to do the least-squares calculation again, within the absorption bands, with the data points spaced 2 cm^{-1} apart. The author undertook to code these finely spaced data for punching onto cards. Although these data were probably extended away from the bands farther than necessary, a set of data decks now exists for plume gases and their mixtures which can support studies of this kind.

Another phase of the work, undertaken by the author, was to prepare a preliminary program for the IBM 370 which would calculate the band integrals and their ratios which were required to get mole fractions from the data decks. Three versions of this program were written and the last is most convenient to use and its output is the easiest to interpret. A copy of the program is Appendix II. Appendices III and IV are samples of the output of the program and Appendix V contains some hints about how to use the program. These hints are in lieu of "comments" within the program itself.

It should be noted that integration is accomplished with the trapezoid rule. The reason for this is that the data has regions of unevenly spaced points and to do the integrals with Simpson's Rule would require breaking up the integral into regions.

The results of the calculations for the mixtures studied are shown in the table below:

Mixture	Molecular Concentration in Cryodeposited Mixture	Ratio of Band Strength in Mixture to Band Strength in Pure Cryodeposit	Band Used Cm^{-1}
N_2-CO_2	.253 CO_2	.220 CO_2	2260-2500
N_2-CO	.203 CO	.189 CO	2080-2240
N_2-H_2O	.123 H_2O	.0925 H_2O	1490-1720
CO_2-CO	.498 CO_2	.400 CO_2	2260-2500
	.495 CO	1.24 CO	2080-2240
Plume 1	.172 CO_2	.151 CO_2	2260-2500
	.10 CO	.435 CO	2080-2240
	.225 H_2O	.0241 H_2O	1490-1720

For purposes of discussion the results may be divided into two sections. One section consists of the first three mixtures studied which are mixtures of N_2 with three plume gases, CO, CO_2 and H_2O. In these mixtures the hypothesis of a band strength linear with concentration appears to hold. The evidence for this is the agreement shown in the table between the ratio of the band strength in the mixture to the band strength in pure cryodeposit and the molecular concentrations obtained by the chemistry lab. It appears that in these mixtures N_2 is merely a diluent. It in no way changes the strength of the bands although their shapes are frequently altered. The final two mixtures studied are the second group. In both cases they are mixtures which include two or more infrared-active molecules, CO_2 and CO and a simulated plume exhaust mixture with CO, CO_2 and H_2O plus N_2. Here the ratios of the band integrals do not agree with the chemical laboratory's analysis. The hypothesis of a linear dependence of band strength on concentration does not hold for these mixtures. Although this rules out analysis of cryodeposits based on band strengths linear with concentration, it is possible to reach a few "non-trivial" conclusions about the infrared absorption of windows which have mixtures of plume gases cryodeposited on them. If it occurs in combination with CO and CO_2 as it does in a rocket plume, water may not be as strong an absorber as might be supposed. The band integral for the water is less in the plume mixture than would be expected on the basis of the linear hypothesis.

It is interesting to spectulate about the causes of these effects. In a cryodeposit the distance between molecules is very much smaller than in a gas. The energy level structure of a cryodeposit is not just a collection of the

energy levels from its constituent molecules. The electomagnetic fields generated by charges in one molecule may influence neighboring molecules.

Test of the Second Hypothesis

One of the most elegant theorems of radiation theory is the so-called Kramers-Kronig relationship which states that the real part and the imaginary part of the complex index of refraction are not independent; they are related by integral equations. Recently these relations have been put in what is called subtractive form. This form is useful for working with limited parts of the spectrum because the results depend only slightly on the values of the indices outside the region of study. Palmer has written a program to calculate the real part of the index of refraction from the imaginary part. For the most part the results are in good agreement with the indices obtained from the least-squares program. In very strong or very sharp absorption bands, however, the values given by the Kramers-Kornig formalism do not fluctuate as strongly as those given by the least-squares program.

Recommendations

It would seem that it may be possible to establish an analysis technique for cryodeposited gases based on their infrared absorption but the simple hypothesis of a linear relationship between band strength and concentration will not work if more than one infrared molecule is present. It appears that the next step is to consider some simple case, like a half-and-half mixture of CO and CO_2, and see what happens to the infrared-active energy levels as they approach each other.

References

1. Abrenkiel, R. K., "Modified Kramers-Kronig Analysis of Optical Spectra." Journal of the Optical Society of America, 61, 1651-1655 (1971).

2. McCracken, Daniel, A Guide to Fortran-IV Programming, Wiley (1965).

3. Palmer, Kent, An Attempt to Quantify Binary Cryodeposits, Report of Summer Research at AEDC (1976).

4. Palmer, Kent, Subtractive Kramers-Kronig Analysis Program for IBM-370.

5. Pipes, John, et al, Infrared Transmission of Clean and Contaminated Cryo-cooled Optical Windows, AEDC Technical Report (to be published).

6. Existing Computer Program at AEDC, written by Palmer & Roux, Jul 1976.

****

N2-H2O

****

START = 2. STOP = 300. INC = 1. NUO = 2000 CM-1. NO = 1.3030000

NU (CM-1)	ALPHA (CM-1)	N	K	N-LS INDEX	TRANSMITTANCE
3700	446.3575		0.009600		
3690	1743.5085	1.297371	0.037600	1.285000	0.982982
3680	1026.6222	1.316844	0.022200	1.347000	0.981208
3670	470.4095	1.317933	0.010200	1.323000	0.981168
3660	174.7731	1.312632	0.003800	1.320000	0.981722
3650	59.6274	1.308192	0.001300	1.302000	0.982172
3640	50.3157	1.304038	0.001100	1.303000	0.982587
3630	95.7934	1.302438	0.002100	1.303000	0.982745
3620	54.5883	1.300668	0.001200	1.303000	0.982921
3610	54.4375	1.298768	0.001200	1.303000	0.983108
3600	81.4301	1.296528	0.001800	1.297000	0.983327
3590	112.7832	1.295051	0.002500	1.298000	0.983471
3580	143.9603	1.293081	0.003200	1.304000	0.983662
3570	206.3649	1.291545	0.004600	1.294000	0.983810
3560	263.9440	1.289568	0.005900	1.289000	0.983998
3550	370.2681	1.287797	0.008300	1.296000	0.984162
3540	547.1649	1.287048	0.012300	1.305000	0.984219
3530	700.8768	1.288607	0.015800	1.289000	0.984050
3520	769.0051	1.290744	0.017400	1.308000	0.983834
3510	798.3541	1.292782	0.018100	1.309000	0.983632
3500	796.0796	1.294119	0.018100	1.304000	0.983502
3490	785.0337	1.295153	0.017900	1.307000	0.983403
3480	756.5458	1.295237	0.017300	1.311000	0.983398
3470	754.3718	1.294667	0.017300	1.308000	0.983454
3460	804.3734	1.293992	0.018500	1.305000	0.983512
3450	867.0796	1.294081	0.020000	1.312000	0.983492
3440	925.0859	1.294556	0.021400	1.316000	0.983435
3430	987.0507	1.295640	0.022900	1.328000	0.983317
3420	1005.0615	1.296693	0.023400	1.327000	0.983210
3410	1032.7169	1.297553	0.024100	1.321000	0.983119

Appendix II

Band Integration Program

```
LEVEL  21                    MAIN                    DATE = 77208          10/10/42

      DIMENSIONNH(9,20),NL(9,20),FR(9),  N(1500),AK(1500),Q(9,20),QM(20)
      DIMENSION NB(9)
      DIMENSION GS(9,10),TTL(10),DTTL(20)  ,AN(1500)
      WRITE(6,1)
    1 FORMAT(1H1)
      READ(5,11)TTL,NG
   11 FORMAT (10A4,I10      )
      WRITE(6,21)TTL,NG
   21 FORMAT  (10A4,16HNUMBER OF GASES=,I10,5X      //                    )
      DO 80 K=1,NG
      READ (5,31)(GS(K,I),I=1,10),FR(K),NB(K)
   31 FORMAT(      10A4,  F10.5,I10)
      NZ=NB(K)
  300 READ  (5,41)(NH(K,J),NL(K,J), J=1,NZ)
   41 FORMAT(4X,I10,6X,I10)
  200 READ(5,51)DTTL
   51 FORMAT(20A4)
      WRITE (6,151)DTTL
  151  FORMAT(1X,20A4//)
      DO 40 L=1,1500
      READ (5,61)AN(L),AK(L)
   61 FORMAT(F13.0,25X,G12.5)
      N(L)=INT(AN (L) + .000005)
      IF(N(L))7,50,40
   40 NDC=L
    7 WRITE (6,71)
   71 FORMAT (1X, 19HERROR:W.N. NEGATIVE)
      STOP
   50 DO 80 JA=1,NZ
      DO 55 LA=1,NDC
      INH=LA-1
      IF(N(LA)-NH(K,JA)) 55,60,55
   55 CONTINUE
   60 DO 65 LB=1,NDC
      INL=LB
      IF(N(LB)-NL(K,JA))  65,70,65
   65 CONTINUE
   70 Q(K,JA)=0.0
      DO 12 LC=INH,INL
      DN=N(LC)-N(LC+1)
      IF(DN)17,17,12
   17 WRITE (6,101)
  101 FORMAT (26HERROR:CARD OUT OF SEQUENCE)
```

```
      STOP
   12 Q(K,JA)=Q(K,JA)+((AK(LC)+AK(LC+1))/2.0)*DN
      WRITE (6,121) (GS(K,M),  M=1,10),FR(K)    ,NB(K)
  121 FORMAT (1X,10A4,F10.5, 16HNUMBER OF BANDS=,I10)
      WRITE(6,73)Q(K,JA),NL(K,JA),NH(K,JA)
   73 FORMAT(1X,14HBAND INTEGRAL=,E10.5,5X,18HLOWER WAVE NUMBER=, I10,5X
     &,19HHIGHER WAVE NUMBER=,I10///)
   80 CONTINUE
      NC=NG-1
      JC=0
      DO100  KB=1,NC
      NX=NB(KB)
  500 DO 100 JB=1,NX

      JC=JC+1
      BMF = Q(NG,JC)/Q(KB,JB)
      WRITE(6,181) BMF,  (GS(KB,M), M=1,10), FR(KB)
  181 FORMAT (1X,32H BAND MOLE FRACTION FOR MIXTURE=,F10.5,10A4, F10.5/)
      WRITE (6,81)TTL,NL( KB,JB),NH(KB,JB)
   81 FORMAT(1X,10A4,17HLOWER WAVENUMBER=,I10,5X,17HUPPER WAVENUMBER=,
     &I10///)
  100 CONTINUE
      STOP
      END
```

PLUME I NUMBER OF GASES= 4

CO2 (20 DEGREES K)

CO2
INTEGRAL=.60001E 02 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 1
2260 HIGHER WAVE NUMBER= 2500

CO (20 DEGREES K)

CO
INTEGRAL=.21327E 01 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 1
2080 HIGHER WAVE NUMBER= 2240

H2O

H2O
INTEGRAL=.20414E 02 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 1
1490 HIGHER WAVE NUMBER= 1720

PLUME MIXTURE 1 (20K)

PLUME I
INTEGRAL=.90771E 01 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 3
2260 HIGHER WAVE NUMBER= 2500

PLUME I
INTEGRAL=.92721E 00 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 3
2080 HIGHER WAVE NUMBER= 2240

PLUME I
INTEGRAL=.49188E 00 LOWER WAVE NUMBER= 0.0 NUMBER OF BANDS= 3
1490 HIGHER WAVE NUMBER= 1720

Appendix IV

Band Integration Program Output

for Plume Mixture I Mole Fractions

```
MOLE FRACTION FOR MIXTURE= 0.15128    CO2                    0.0                    2500
 PLUME  I                        LOWER WAVENUMBER=  2260    UPPER WAVENUMBER=

MOLE FRACTION FOR MIXTURE= 0.43475    CO                     0.0                    2240
 PLUME  I                        LOWER WAVENUMBER=  2080    UPPER WAVENUMBER=

MOLE FRACTION FOR MIXTURE= 0.02410    H20                    0.0                    1720
 PLUME  I                        LOWER WAVENUMBER=  1490    UPPER WAVENUMBER=
```

Appendix V

Use of the Band Integration Program

I. Cards Read

A. Title card with gas mixture and temperature of the deposit in columns 1-40, and the number of gases (IR-active) in the deposit. Include the mixture but not gases with no absorption bands like N_2.

B. For each IR-active gas there should be a gas card with the name of the gas in columns 1-40, the fraction in shich that molecule is present in the mixture, if known, and the number of bands for the gas which are to be integrated in columns 51-60 (right justified). The fraction is optional and may not be known. It is simply printed in the output for comparison. If it is left blank you will get 0.0.

C. For each band specified above an infrared band card must be in the data. The higher limit of the band should appear right justified in columns 5-14 and the lower limit of the band should appear right justified in columns 21-30.

D. Each of the data decks has a title, that card is read and written in the output. You may use all columns of the card for the title.

E. For each of the gases up to 1500 data cards may be read. The format statement is a compromise. We want an integer in the first field but some of the data cards have decimal points in them and will not be read unless we use F-format. The format in the program reads all of the data cards we have. The last card of each constituent deck and the mixture deck should be a blank card. This will be a signal that the data deck for that particular gas has been completely read.

F. The total data deck is arranged with the package for each constituent gas (arranged as stated above) followed by the package for the mixture. The band cards in the mixture package must be in the same order as they are encountered in the constituent decks because the program finds corresponding bands by counting.

II. The output consists of the band integrals for the infrared-active constituents and for the mixture. This is followed by the mole fractions calculated from the band strengths. If the actual fractions are known and entered into the cards they will appear in the output on the same line as the computed mole fractions.

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

Sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Conducted by

AUBURN UNIVERSITY AND OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

ROCKET EXHAUST GAS DYNAMICS

Prepared by:	Peter C. Sukanek, Ph.D.
Academic Rank:	Assistant Professor
Department and University:	Department of Chemical Engineering Clarkson College
Assignment:	Arnold Air Force Station Arnold Engineering and Development Center Engine Test Facility Technology Applications Branch
USAF Research Colleague:	Robert P. Rhodes
Date:	August 8, 1977
Contract No.	F44620-75-C-0031

ABSTRACT

Three problems relevant to the predictions of the fluid mechanics of low altitude plumes have been investigated. These include a simple approximation to the plume expansion, an investigation of constant pressure turbulent mixing and the development of a "complete" plume model.

A simple method for calculating constant pressure properties of plumes was extended to include real gas effects. For conditions of interest, these are essentially negligible. The perfect gas approximation is adequate.

Four turbulent viscosity models are investigated. None gives predictions in line with experimental data. While the species distribution and pitot pressures are in many cases not vastly different, the predicted temperatures would give rise to very different values of total radiation. Additional work on compressibility effects and numerical error should be undertaken.

Finally, general equations describing a multi-zone plume capable of accounting for shocks, radial pressure variations and turbulent mixing are developed. The equations are specialized to the case of a three-zone plume and reduced to a set which can be easily solved numerically.

NOMENCLATURE

A_i	area at location i
a	diffusion parameter given by equation (14)
a^1, a_2	constants in 1-equation TKE model
c	mixing parameter
C_D	drag coefficient
C_F	thrust coefficient
$C_\mu, C_{\varepsilon 1}, C_{\varepsilon 2}$	constants in k-ε model
h	static enthalpy
h^o	heat of formation
H	total enthalpy
k	turbulence kinetic energy
ℓ	mixing length
Le	turbulent Lewis number
$\dot{m}$	mass flow rate
M	Mach number
P	pressure
P_r	turbulent Prandtl number
r	radial position, radius
R_α	production rate of species α
S	entropy
T	temperature
u	axial velocity
v	radial velocity
γ	ratio of specific tests
ε	turbulence energy dissipation
Ψ	stream function
ρ	density
$\sigma_k, \sigma_\varepsilon$	"Prantl numbers" for k and ε
$\tau_{xr}, \tau_{rr}, \tau_{oo}$	turbulent shear and normal stresses
μ	turbulent viscosity
ω	mass fraction
R	gas constant

SUBSCRIPTS

a	ambient
c	chamber or core
e	exit plane
I	plume-freestream interface
m	mixing region
o	centerline
t	throat
∞	freestream
α	species

INTRODUCTION

The prediction of the fluid mechanical structure of low altitude rocket exhaust is of interest to the Air Force in a variety of applications. The ability to predict infrared or ultraviolet radiation from, secondary condensation in or microwave or laser interference with the exhaust plume all require a knowledge of the temperature, pressure and chemical species distribution in the plume as a starting point. The purpose of the research described below is to examine various aspects of the overall problem. These deal with the development of a simple model for predicting the start-line properties of low altitude ($\lesssim$ 60 km) plumes, the evaluation of turbulence models and numerical formulations for constant pressure plumes and the investigation of a multi-zone, one-dimensional underexpanded plume model.

The three problem areas are distinct, and so are treated separately in the sections which follow. However, they are, in a sense, all interrelated. The calculation of the plume flow field at low altitude requires a model for turbulence. At the present time, no turbulence model is clearly superior to any other. Yet the choice of the model greatly influences what the predicted flow field is, as will be illustrated below. The only way to make this choice is to compare the prediction of several of the most reasonable models to available data.

Unfortunately, the task of making predictions is further complicated since the equations must be solved numerically. This introduces another uncertainty into the analysis. As will be shown below, the numerical "constitutive equations", i.e., the manner in which the partial derivatives in the governing equations are approximated by a finite-difference scheme, may have as much effect on the predictions as the turbulence model employed in the calculation. This adds an additional dimension of uncertainty to the plume flow field predictions, an uncertainty that must be taken into account when the flow field is used as input to further calculations.

The difficulties involved in calculating plume gas dynamics goes beyond the items just mentioned. If the exhaust gas left the rocket nozzle in a parallel stream at a pressure equal to the local static pressure, the mixing and combustion with the freestream could, in principle, be calculated from a knowledge of the appropriate turbulence model. However, in flight, the nozzle exit pressure is usually greater than the freestream static pressure. The initial portion of the exhaust plume is dominated by a series of expansions and compressions, and the resulting shock waves. This increased the complexity of the calculation tremendously. The pressure in the viscous mixing layer is now a function of both axial and radial distance. The shock waves, generated in the inviscid core, propagate into the shear layer where they interact with

the turbulent flow and are dissipated. This flow field cannot be calculated at the present time. A second topic of the present investigation deals with two ways in which this complex flow field can be approximated.

The first technique treats this initial portion of the plume as a lumped region in which the plume expands to ambient pressure. The expansion is presumed to be one-dimensional, so that the constant pressure startline is governed by a single temperature. The effects of a variable heat capacity of plume gas on the startline properties is investigated by comparing the expansion of a perfect gas to that of a real one.

A second type of model for this expansion is also presented. The plume is assumed to consist of a number of discrete zones, each of which is perfectly mixed at any axial station. This model is capable of predicting the axial distance at which the plume can be considered to be at constant pressure and give some indication of the variation of properties with radial position at that point. In addition, the model can give some insight into the effects of the viscous-inviscid coupling on the plume.

This paper is divided into three sections, one dealing with each of the three topics discussed above. In the following section, the one-dimensional expansion to ambient model is discussed. A comparison between the real gas and ideal gas expansions are presented. The following section treats the constant pressure mixing problem. The formulation of the finite-difference equations is discussed, and two alternate approaches are described. A number of turbulent mixing models are also presented. The effect of these models on selected plume properties, and how the predictions compare with measurements, is shown. In the final section, the equations governing a multi-zone plume model are presented and a solution technique is discussed.

SIMPLE EXPANSION MODEL

The basic componenets of this model are described elsewhere (Ref. 1). Only the major features are reported here.

Figure 1 shows a plume expansion from a rocket nozzle. The inviscid core and the shock waves generated in it are surrounded by a mixing layer which grows until it eventually reaches the axis. Somewhere in the initial region of the plume, the pressure reaches the local static pressure. In the simple expansion

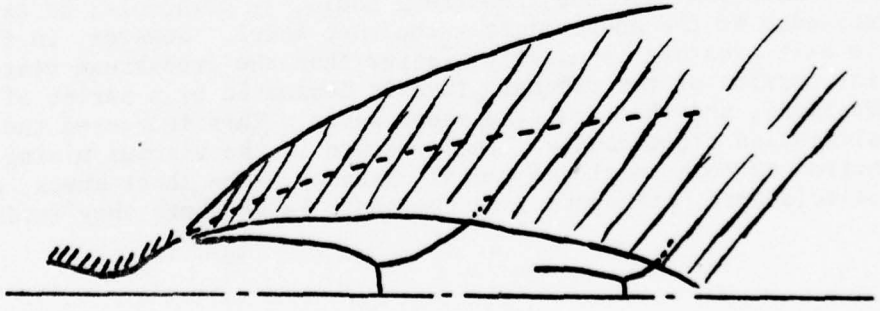


Figure 1. Plume Expansion

model, we ignore the details of the expansion process and assume only that somewhere downstream of the actual exit plane the matched pressure point is realized. With this assumption, we can then write the following equations for conservation of mass, energy and momentum:

$$\rho_e u_e r_e^2 = \rho u r^2$$

$$h_e + \tfrac{1}{2} u_e^2 = h + \tfrac{1}{2} u^2$$

$$\rho_e u_e^2 r_e^2 - \rho u^2 r^2 + (P_e - P_\infty) r_e^2 + \tfrac{1}{2} \rho_\infty u_\infty^2 r^2 C_D = 0 \qquad (1)$$

where C_D is the drag coefficient, defined as:

$$C_D \equiv \frac{1}{\frac{1}{2}\rho_\infty u_\infty^2 \pi r^2} \int (P_I - P_\infty)\, 2\pi r\, dr. \qquad (2)$$

In writing the mass balance, equation $(1)_1$, we have assumed that no freestream gas is entrained into the plume.

The drag coefficient, C_D, is a critical parameter in these equations. It can be shown that it represents the entropy production in the expansion process (Ref. 1). In the actual expansion entropy is produced in two ways: first, by the shock waves in the inviscid core, and second, by the turbulent mixing which takes place in the shear layer. The inclusion of only one of these two mechanisms in the evaluation of C_D would lead to an underestimate of the entropy change.

In the present model for the plume expansion, an empirical evaluation of the drag is used based on the "Universal Plume" model of Jarvinen and Hill (Ref. 2). The plume drag is assumed to be given by (Ref. 1):

$$C_D = \frac{16}{9\pi} \left[\frac{C_{F MAX}}{C_F} - 1 \right]^{1/2} \qquad (3)$$

where the thrust coefficients, C_F and $C_{F_{Max}}$, are:

$$C_F = C_{F MAX} \left[1 - \left(\frac{P_e}{P_c} \right)^{\frac{\gamma-1}{\gamma}} \right]^{1/2} + \frac{P_e - P_\infty}{P_c} \frac{A_e}{A_c}$$

$$C_{F MAX} = \left\{ \frac{2\gamma^2}{\gamma-1} \left(\frac{2}{\gamma+1} \right)^{\frac{\gamma+1}{\gamma-1}} \right\}^{1/2}. \qquad (4)$$

For the special case of a perfect gas (i.e., constant γ), equations (1) can be solved in closed form to give the conditions at the matched pressure point:

$$u = (1+f)\,u_e$$

$$T = \left[1 - m_e^2(\gamma-1)f - \tfrac{1}{2}m_e^2(\gamma-1)f^2\right]T_e$$

$$r = \left(\frac{P_e}{P_\infty}\,\frac{T}{T_e}\,\frac{u_e}{u}\right)^{1/2} r_e \tag{5}$$

where

$$f = \frac{-B + \sqrt{B^2 + 4AC}}{2A} \tag{6}$$

and

$$A = \gamma m_e^2 + \frac{\gamma-1}{4}\,m_e^2\,\gamma_\infty\,m_\infty^2\,C_D$$

$$B = \gamma m_e^2 + \frac{\gamma-1}{2}\,m_e^2\,\gamma_\infty\,m_\infty^2\,C_D - 1 + \frac{P_\infty}{P_e}$$

$$C = 1 - \frac{P_\infty}{P_e} + \tfrac{1}{2}\gamma_\infty\,m_\infty^2\,C_D\,. \tag{7}$$

For a "real gas" (variable γ) plume, no closed form solution is possible. The composition of the gas and the variation of the thermodynamic properties of the gas with temperature must be known. The real gas expasnion problem has been solved numerically using a computer program which calculates the expansion ratio and temperature of a gaseous mixture expanded from known conditions to a specified pressure and entropy. For the problem at hand, the known conditions are those at the exit plane and the specified pressure is ambient static. The entropy change is unknown, but must be such that the momentum balance, equation $(1)_3$, is satisfied. Hence, an initial value of ΔS is guessed and the error in the momentum balance is used to compute a revised value for ΔS. The secant method for estimating the revised value was used (Ref. 3):

$$\Delta S_{n+1} = \frac{\Delta S_{n-1}\,\delta_n - \Delta S_n\,\delta_{n-1}}{\delta_n - \delta_{n-1}} \tag{8}$$

where

$$\delta = \rho u^2 (r/r_e)^2 - \rho_e u_e^2 - (P_e - P_\infty) - \tfrac{1}{2}\gamma_\infty\,m_\infty^2\,P_\infty\,C_D\,(r/r_e)^2. \tag{9}$$

The iteration scheme is remarkably well-behaved, requiring only about 10 iterations to converge to a solution.

Figure 2 shows a comparison between the predicted perfect gas and ideal gas temperatures at the matched-pressure point for the plume and freestream properties given in Table 1. As is apparent from the figure, the difference between the two sets of predictions increased as the altitude increases. However, the difference between the two types of prediction is almost negligible throughout the altitude region considered and is probably within the errors of the other approximations.

One aspect of this technique must be mentioned. When the global balance model is employed to calculated the fictious matched-pressure location, all information concerning the axial distance from the nozzle exit at which this constant pressure location occurs is lost. We could assume that the calculation should be started at the real exit plane, since the constant pressure mixing models which are used downstream of this point start with a zero or near zero thickness of the shear layer, and it is at the real exit plane that the actual shear layer starts to grow. On the other hand, we might expect that the actual constant (or more correctly, nearly-constant) pressure location is at some finite distance from the exit plane. If our drag coefficient were correct, indicating a correct amount of entropy increase, then we have implicitly assumed some mixing takes place during the expansion. Unfortunately, our assumption of uniform properties at this point is in conflict with this line of reasoning. If a constant pressure mixing model is used downstream of our calculated location, the mixing model must, of necessity, start with a "zero" thickness shear layer.

I believe that discussions of this sort are pointless. By employing this approximation, we have abrogated our right to any knowledge concerning at what axial location to start the mixing calculation. It seems most reasonable to assume the mixing starts at the actual exit plane. However, there is no a priori reason for the predicted spatial distribution of plume properties to bear any resemblance to reality. Indeed, there must be disagreement between measurements and predictions in some region of the plume regardless of the axial location selected to start a constant pressure mixing calculation.

---

[1] Convergence is defined as a value of ΔS such that

$$\delta \leq 0.00001 \left[\rho_e u_e^2 + (P_e - P_\infty) \right]$$

TABLE 1. EXIT PLANE AND FREESTREAM PROPERTIES

Exit Plane: $T = 1854°K$, $P = 0.924$ atm, $u = 9104$ ft/sec, $\gamma = 1.23$, $\varepsilon = 6$, $M = 3.06$, $P_c = 57.76$ atm

Species	Mole Fraction
H	1.62 (-3)[1]
O	1.51 (-5)
OH	8.01 (-4)
NO	9.71 (-2)
CO	0.0528
H_2	0.0659
O_2	5.15 (-5)
H_2O	0.441
CO_2	0.0784
N_2	0.349

Freestream: $\gamma = 1.4$

Altitude (km)	Pressure (atm)	Mach Number
5	0.531	0.687
15	0.119	1.76
25	0.0248	2.85
35	5.65 (-3)	3.73
50	7.89 (-4)	4.94
59	2.42 (-4)	5.91

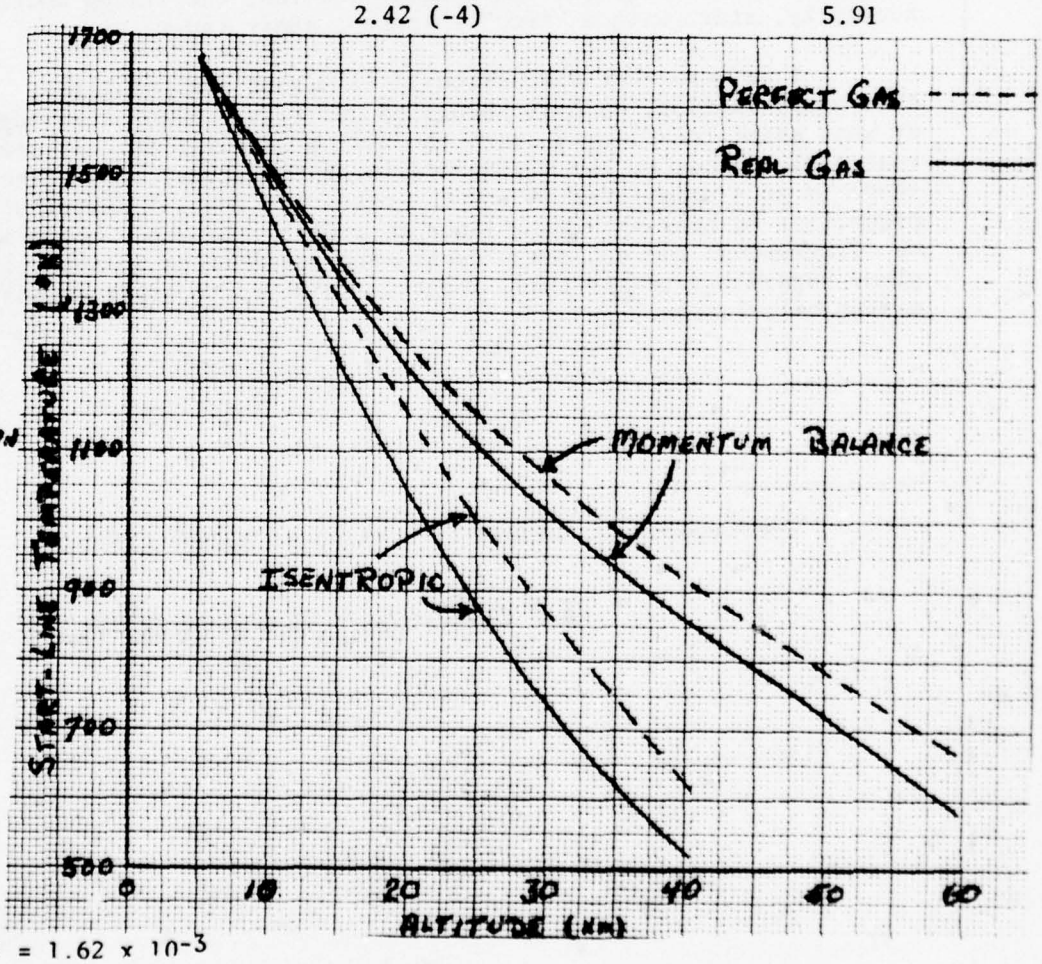


FIGURE 2. COMPARISON OF REAL AND PERFECT GAS START-LINE TEMPERATURES.

[1] $1.62\ (-3) = 1.62 \times 10^{-3}$

EVALUATION OF TURBULENCE MODELS

Whether or not a simple procedure such as the one just described is used to begin a plume calculation, the plume model must contain a mathematical description of the turbulence. In the case of a constant pressure plume, the problem is certainly much simpler, since only the axial momentum equation, containing only one stress component, needs to be solved. In this section, discussion will be limited to that particular case, constant pressure mixing. However, one would hope that a turbulence model valid in this case would also be valid in the variable-pressure situation as well.

The turbulence model selected for incorporation into the plume calculation should be the most "correct" one. At the present time, there are few theoretical guidelines for the form of such a model. The ultimate test of "correctness" is the ability of the model to reproduce available mixing data. For such an important subject, it is surprising that so few comparisons have been performed. Indeed, for the most part, what has been compared in the past to experimental data is both the turbulence model and the numerical framework in which the turbulence model sits. To evaluate his model, the turbulence modeler typically writes his own finite-difference program. Hence, there are almost as many computer programs as there are turbulence models. The comparisons performed thus far are really comparisons of the total package, turbulence and numerical models. In this section, we will provide some comparisons between several different turbulence models all used in the same numerical scheme.

Prior to presenting the comparisons, we will review the equations, both conservation and constitutive, and the numerical procedure used in the program. Low Altitude Plume Program (Ref. 4) was employed, with the modifications discussed below. For an axisymmetric turbulent flow, the equations of conservation of mass, momentum, energy and species mass are written in boundary layer form as[1]:

$$\frac{\partial}{\partial x}(\rho u) + \frac{1}{r}\frac{\partial}{\partial r}(\rho v r) = 0$$

$$\rho u \frac{\partial u}{\partial x} + \rho v \frac{\partial u}{\partial r} = -\frac{dp}{dx} + \frac{1}{r}\frac{\partial}{\partial r}\left(r\mu \frac{\partial u}{\partial r}\right)$$

$$\rho u \frac{\partial H}{\partial x} + \rho v \frac{\partial H}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\left\{\frac{\mu r}{\rho_r}\left[\frac{\partial H}{\partial r} + \left(\frac{P_r}{\sigma_k} - 1\right)\frac{\partial k}{\partial r} + \left(\frac{P_r - 1}{2}\right)\frac{\partial u^2}{\partial r} + (Le - 1)\sum_\alpha h_\alpha \frac{\partial \omega_\alpha}{\partial r}\right]\right\}$$

$$\rho u \frac{\partial \omega_\alpha}{\partial x} + \rho v \frac{\partial \omega_\alpha}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{Le}{P_r}\mu r \frac{\partial \omega_\alpha}{\partial r}\right) + R_\alpha \qquad (10)$$

---

[1] In writing the energy equation in this form, we have ignored terms such as $\Sigma \overline{\nabla \omega_\alpha' h_\alpha'}$ and $\Sigma \overline{u' h_\alpha \omega_\alpha'}$ which appear after the time averaging. This is standard practice, although it has not been justified.

These are supplemented by conservation equations for turbulence kinetic energy and turbulent energy dissipation:

$$\rho u \frac{\partial k}{\partial x} + \rho v \frac{\partial k}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r\mu}{\sigma_k}\frac{\partial k}{\partial r}\right) + \mu\left(\frac{\partial u}{\partial r}\right)^2 - \begin{cases} a_2 \rho k^{3/2}/\ell \\ \rho\varepsilon \end{cases}$$

$$\rho u \frac{\partial \varepsilon}{\partial x} + \rho v \frac{\partial \varepsilon}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r\mu}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial r}\right) + C_{\varepsilon 1}\frac{\varepsilon}{k}\mu\left(\frac{\partial u}{\partial r}\right)^2 - C_{\varepsilon 2}\rho\varepsilon^2/k \qquad (11)$$

These parameters are used in the definition of effective turbulent viscosity by some modelers. The form of the energy dissipation term in equation (11), depends on whether or not ε is required in the definition of μ.

The conservation equations as they stand are valid, within the spirit of the approximations discussed above, for constant density flows only. The effects of compressibility must be accounted for elsewhere. The practice has been to incorporate a compressibility correction into the equation for the turbulent viscosity or in the dissipation. These corrections will be discussed below.

Introducing a stream function, Ψ, defined as:

$$\psi \frac{\partial \psi}{\partial r} \equiv \rho u r$$

$$\psi \frac{\partial \psi}{\partial x} \equiv -\rho v r \qquad (12)$$

and transforming equations (10) and (11) from (x,r) coordinates to (x,Ψ) coordinates allows us to write the equations as:

$$\frac{\partial F}{\partial x} = \frac{1}{\psi}\frac{\partial}{\partial \psi}\left(\frac{a}{b}\frac{\partial F}{\partial \psi}\right) + P - D \qquad (13)$$

where

$$a \equiv \frac{\mu \rho u r^2}{\psi} \qquad (14)$$

and the function F, P and D and the parameters b are given in Table 2.

Table 2. Functions for Equation (13)

F	b	P	D
u	1	o	$\frac{1}{\rho u}\frac{dp}{dx}$
H	Pr	$\frac{1}{\psi}\frac{\partial}{\partial\psi}\left\{\frac{a}{\sigma_r}\left[\left(\frac{\sigma_r}{\sigma_k}-1\right)\frac{\partial k}{\partial\psi}+\frac{Pr-1}{2}\frac{\partial u^2}{\partial r}\right.\right.$ $\left.\left.+(Le-1)\sum h_\alpha\frac{\partial\omega_\alpha}{\partial\psi}\right]\right\}$	o
ω_α	Pr/Le	$R_\alpha/\rho u$	o
k	σ_k	$\frac{a}{\psi}\left(\frac{\partial u}{\partial\psi}\right)^2$	$\begin{cases} a_2 k^{3/2}/lu \\ \varepsilon/u \end{cases}$
ε	σ_ε	$C_{\varepsilon 1}\frac{\varepsilon}{k}\frac{a}{\psi}\left(\frac{\partial u}{\partial\psi}\right)^2$	$C_{\varepsilon 2}\varepsilon^2/uk$

Equation (13) must be solved numerically using an appropriate finite difference technique. The one used by LAPP is similar to that employed by Patankar and Spalding (Ref. 5). It is obtained by multiplying equation (13) by $\Psi\, d\Psi$ and integrated across each stream tube, from $\Psi-$ to $\Psi+$. (The exact meaning of $\Psi-$ and $\Psi+$ will become more apparent when we discuss possible choices for stream tube positions and the different numerical formulations. These will be discussed shortly.) Some of the P and D functions cannot be integrated exactly. In these cases, P and D are replaced by "average values" over the stream. The result is an equation of the form:

$$\frac{\partial F}{\partial x}\frac{\psi_+^2-\psi_-^2}{2}=\frac{a}{b}\frac{\partial F}{\partial\psi}\Big|_+-\frac{a}{b}\frac{\partial F}{\partial\psi}\Big|_-+\bar{P}\frac{\psi_+^2-\psi_-^2}{2}-\bar{D}\frac{\psi_+^2-\psi_-^2}{2} \qquad (15)$$

The derivatives are then approximated using finite differences as:

$$\frac{\partial F}{\partial x}\cong\frac{F_{n+1,m}-F_{n,m}}{\Delta x}$$

$$\frac{a}{b}\frac{\partial F}{\partial\psi}\Big|_+\cong\left(\frac{a}{b}\right)_+\frac{F_{n,m+1}-F_{n,m}}{\Delta\psi}$$

$$=\frac{a_{n,m+1}+a_{n,m}}{2b}\frac{F_{n,m+1}-F_{n,m}}{\Delta\psi} \qquad (16)$$

$$\frac{a}{b}\frac{\partial F}{\partial\psi}\Big|_-\cong\frac{a_{n,m}+a_{n,m-1}}{2b}\frac{F_{n,m}-F_{n,m-1}}{\Delta\psi} \qquad (17)$$

The average values of P and D are taken as the values at the center of the streamtube[1]. The LAPP program uses "implicit differences" in the evaluation of the pressure gradient and the chemical production terms. In this formalism, the terms are evaluated at the downstream axial station. The result is a set of finite difference equations of the form:

$$F_{n+1,m} - \frac{\Delta x}{\rho_{n,m} u_{n,m}} P^{+}_{n+1,m} = F_{n,m} + \frac{\Delta x}{2\psi_m \Delta\psi^2 b}\left\{(a_{n,m+1}+a_{n,m})(F_{n,m+1}-F_{n,m})\right.$$

$$\left.-(a_{n,m}+a_{n,m-1})(F_{n,m}-F_{n,m-1})\right\} + \frac{\Delta x}{4\psi_m \Delta\psi^2} P - \frac{\Delta x}{u_{n,m}} D_{n,m} \qquad (18)$$

The appropriate values of F, P^{+}, b, P and D are given in Table 3.

Table 3. Definitions for Eqatuion (18)

F	P^{+}	b	P	D
u	$-\frac{dP}{dx}$	1	o	o
H	o	P_r	$\frac{2}{P_r}\{(a_{n,m+1}+a_{n,m})[(\frac{P_r}{\sigma_k}-1)$ $(k_{n,m+1}-k_{n,m})+\frac{P_r-1}{2}(u^2_{n,m+1}-u^2_{n,m})+(Le-1)\Sigma h_\alpha(\omega_{\alpha_{n,m+1}}-\omega_{\alpha_{n,m}})]$ $-(a_{n,m-1}+a_{n,m})[(\frac{P_r}{\sigma_k}-1)(k_{n,m}-k_{n,m-1})+\frac{P_r-1}{2}(u^2_{n,m}-u^2_{n,m-1})$ $+(Le-1)\Sigma h_\alpha(\omega_{\alpha_{n,m}}-\omega_{\alpha_{n,m-1}})]\}$	o
ω_α	R_α	P_r/Le	o	o
k	o_o	σ_k	$a_{n,m}(u_{n,m+1}-u_{n,m-1})^2$	$\{a_2 k^{3/2}/l$
ε	o	σ_ε	$C_{\varepsilon 1}\frac{\varepsilon_{n,m} a_{n,m}}{k_{n,m}}(u_{n,m+1}-u_{n,m-1})^2$	$C_{\varepsilon 2}\varepsilon^2/k$

The finite difference equations take on a different form at the centerline of the flow and at the edge. In the latter case, we can define both the N and N-1 streamtube as belonging to the freestream. In this case, equation (18) still holds provided we realize:

$$F_{n,N} = F_{n,N-1} \qquad (19)$$

in the computer program implimenting these equations.[2]

At the centerline, another form of the finite difference equations must be used to insure the appropriate boundary conditions are satisfied. In the original program (Ref. 4) the centerline is treated by using a limiting form of the governing equations. From (13):

$$\lim_{\psi\to 0}\frac{1}{\psi}\frac{\partial}{\partial\psi}\left(\frac{a}{b}\frac{\partial F}{\partial\psi}\right) = \frac{1}{b}\lim_{\psi\to 0}\left[\frac{a}{\psi}\frac{\partial^2 F}{\partial\psi^2}+\frac{\partial F}{\partial\psi}\frac{\partial a}{\partial\psi}\frac{1}{\psi}\right] \qquad (20)$$

---

[1] Patankar and Spalding (Ref. 5) suggest an alternate definition.

[2] see next page

Using the definition of a, equation (14)

$$\lim_{\psi \to 0} \frac{a}{\psi} = \lim_{\psi \to 0} \frac{\rho \mu u r^2}{\psi} = \lim_{\psi \to 0} 2\rho u \mu r \frac{\partial r}{\partial \psi} = 2\mu$$

$$\lim_{\psi \to 0} \frac{\partial a}{\partial \psi} = \lim_{\psi \to 0} 2\mu \psi \tag{21}$$

Hence, at the centerline, the conservation equation takes the form:

$$\frac{\partial F}{\partial x} = \frac{2\mu_0}{b} \frac{\partial^2 F}{\partial \psi^2} + (P-D)_{\psi=0} \tag{22}$$

In this type of approximation, the centerline of the flow is treated as a disctoninuity. Any of the dependent variables could be represented by a plot like Fig. 3. The mesh points at which the calculations are carried out correspond to the ends of the intervals.

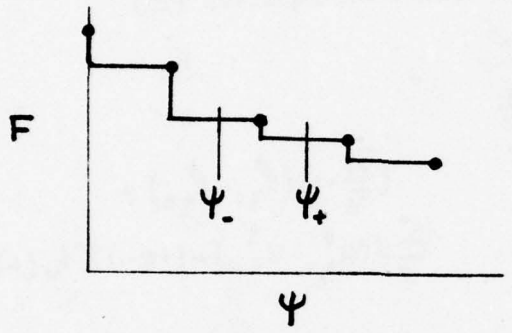


Fig. 3. Mesh Point Distribution in the Original Program

An alternates procedure is to formulate the centerline in a similar fashion as the remainder of the flow. Here, the mesh points can be thought of as distributed as in Fig. 4. The

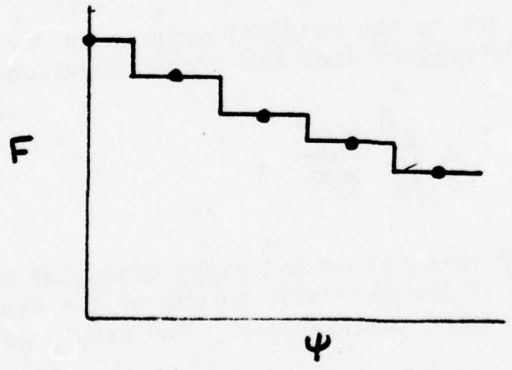


Fig. 4. Alternate Mesh Point Distribution

---

[2] The original Low Altitude Plume Program did not consurve mass, momentum and energy because this condition was not fulfilled.

dependent variables are defined on the center of each streamtube, and every tube contains a finite mass flow. The finite difference equations on the centerline are determined by multiplying equation (13) by $\Psi\, d\Psi$ and integrating from o to Ψ_2.

$$F_{n+1,o} - \frac{\Delta x}{\rho_{n,o} u_{n,o}} P^+_{n+1,o} = F_{n,o} + \frac{4\Delta x}{\Delta\psi^3 b}(a_{n+1}+a_{n,o})(F_{n,1}-F_{n,o}) + \frac{4\Delta x}{\Delta\psi^3}\frac{a_{n,1}+a_{n,o}}{b} P - \frac{\Delta x}{u_{n,o}} D_{n,o} \qquad (23)$$

The values of D and P are defined for each value of F in Table 4.

Table 4. Definitions for Equation (23)

F	b	P^+	P	D
u	1	$-\frac{dp}{dx}$	o	o
H	Pr	o	$(\frac{Pr}{\sigma_k}-1)(k_{n,1}-k_{n,o}) + \frac{Pr-1}{2}(u^2_{n,1}-u^2_{n,o}) + (Le-1)\sum h_\alpha(\omega_{\alpha n,1}-\omega_{\alpha n,o})$	o
ω_α	Pr/Le	R_α	o	o
k	σ_k	o	o	$\begin{cases} a_2 k^{3/2}/l \\ \varepsilon \end{cases}$
ε	σ_ε	o	o	$C_{\varepsilon 2}\varepsilon^2/k$

An additional modification to the original program is the manner in which the radial position is determined from the stream function. From equation (12),

$$\int_o^r \frac{dr^2}{2} = \frac{1}{2}\int_o^\psi \frac{d\psi^2}{\rho u} \qquad (24)$$

The original program evaluates the integral on the right hand side using Simpson's rule. However, because of the step-wise nature of the distribution (Fig. 4), the integral can be evaluated analytically. The result is:

$$r_n = \frac{r_{n+1/2} + r_{n-1/2}}{2} \qquad (25)$$

where

$$r_{n+1/2} = \left[r_{n-1/2}^2 + \frac{2 n \Delta \psi^2}{\rho_n u_n} \right]^{1/2} \tag{26}$$

and

$$r_{1/2} = \frac{\Delta \psi}{2(\rho_0 u_0)^{1/2}} \tag{27}$$

It is easy to show that given a distribution of Ψ, ρ and u both the original and revised procedures predict the same values of r.

The finite difference equations for the constant pressure mixing must now be supplemented with an appropriate expression for the turbulent viscosity, μ. Several different types of constitutive equations for μ have been developed. In this work, three general types were investigated. The first is an algebraic model due to Prandtl (Ref. 6):

$$\mu = \rho c b (U_0 - U_\infty) \tag{28}$$

The second is a turbulent kinetic energy model with an algebraic length scale used by Harsha (Ref. 7):

$$\mu = a_1 \rho k / \left| \frac{\partial u}{\partial r} \right| \tag{29}$$

The third model, developed by Spalding and coworkers (Ref. 8) is a two-equation turbulent kinetic energy model with both the velocity and the length scale given by a differential equation:

$$\mu = C_\mu \rho k^2 / \varepsilon \tag{30}$$

While the length scale does not appear explicitly in equation (29), it is used in the differential equation for k, equation $(11)_1$.

The calculation of plume gas dynamics requires that some compressibility correction to equation (28) to (30) be made. A number of such corrections have been developed in the past. Donaldson and Gray (Ref. 9) arrived at an empirical correction to equation (28) based on experimental measurements of a jet into still air. The constant c in equation (28) is taken to be a function of the Mach number at the half radius. Smoot (Ref. 10) has extended the correlation of Donaldson and Gray to include measurements of jets with external flow. Ting and Libby (Ref. 11) proposed a "compressibility transformation" in which c at any radial position depends on the value of the density ratio integrated from the centerline to that position. Harsha (Ref. 7) assumes that the coefficient a_1 is independent of the Mach number, but assumes that a_2, the coefficient in the dissipation term of equation $(11)_1$ is a function of the turbulence Reynolds number as well as the density ratio. Finally, Dash, et al (Ref. 12) modified equation (30)

using empirical data on the spreading rates of compressible jets. They multiply the turbulent viscosity by a coefficient which depends on the turbulence Mach number.

These various models, both for the numerical procedure and the turbulent viscosity, were used in a comparison between model predictions and experimental data obtained on a rocket plume in supersonic flow. The nozzle was designed to give parallel flow and matched pressure at the exit plane, although some shock waves were observed in the flow field. The predicted nozzle exit properties as well as the values of the turbulence parameters used in the calculation are given in Table 5.

CHEMICAL REACTIONS

$$H + H + M \rightleftarrows H_2 + M$$
$$OH + H + M \rightleftarrows H_2O + M$$
$$H_2 + OH \rightleftarrows H_2O + H$$
$$O + H_2 \rightleftarrows OH + H$$
$$H + O_2 \rightleftarrows OH + O$$
$$CO + OH \rightleftarrows CO_2 + H$$
$$OH + OH \rightleftarrows H_2O + O$$
$$H_2 + O_2 \rightleftarrows OH + H_2$$
$$O + O + M \rightleftarrows O_2 + M$$
$$O + H + M \rightleftarrows OH + M$$
$$CO + O + M \rightleftarrows CO_2 + M$$
$$CO_2 + O \rightleftarrows CO + O_2$$

Table 5. Input Parameters for Plume Calculations

$P = 0.285$ atm, $R_j = 0.14$ ft, $L_e = 1.0$, $P_r = 0.85$

	Jet	Freestream
T(°K)	1294	223
u(ft/sec)	9279	1980
Mole Fractions H_2O	0.366	
N_2	0.316	0.7905
O_2	7.2 (-4)	0.2095
CO	0.12	-
CO_2	0.048	-
H_2	0.126	-
H	0.0116	-
OH	0.0111	-
O	6.17 (-4)	-

Table 5 - Continued

k-ε Model (ER.30)

$$\mu = C(m_t)\, C_\mu\, \rho\, k^2/\varepsilon$$

$$C(m_t) = \begin{cases} 1 & m_t \le 0.03 \\ 0.63 + 0.37 \cos\left[\pi \dfrac{m_t - 0.3}{.334}\right] & 0.03 < m_t \le 0.4 \\ 0.26 & m_t > 0.4 \end{cases}$$

where

$$m_t = \max_{0 \le r \le R} \left(\frac{k}{\gamma R T}\right)^{1/2}$$

and the remaining constants are there given in Ref. 8

k Model (Eq. 29): As in Ref. 7.

The Donaldson-Gray and Ting-Libby models appeared in the original program and were unchanged. The two additional turbulence models, with the appropriate compressibility corrections, were added to the computer code. Initial levels of k and ε were estimated by calculating the turbulent viscosity based on the Donaldson-Gray model and using equations (29) and (30) to find the corresponding values of k and ε[1]. In addition, the Donaldson Gray model was used until the shear layer contained a sufficient number of points (typically 6) to permit the definition of derivatives in this region. This usually occurred at an X/D of about 1.

Consider first the problem of errors introduced by the numerical formulation of the governing differential equations. The two centerline and integration procedures were examined with each of the turbulence models. With the exception of the k-ε turbulence model, essentially identical results were obtained with both numerical techniques. This is not surprising since the revised centerline method uses an average viscosity in the diffusion term, whereas the original method uses the centerline viscosity only. Since the Donaldson-Gray, Ting-Libby and one-equation TKE models all predict essentially the same viscosity at the centerline and the next mesh point, the average of these coincides with the centerline value.

---

[1]Since μ at the exit plane and $\frac{\partial u}{\partial r}$ in the initial region are typically small values, k was chosen as

$$k = \max \begin{cases} \mu \left|\frac{\partial u}{\partial r}\right| / 0.3\rho \\ 0.0001\, u^2 \end{cases}$$

Levels of k in the freestream were chosen as a factor of 100 lower.

The k-ε model does show some differences between the numerical formulations as illustrated in Fig. 5, for the centerline decay and a radial profile at one axial station. In this model, the viscosity at a

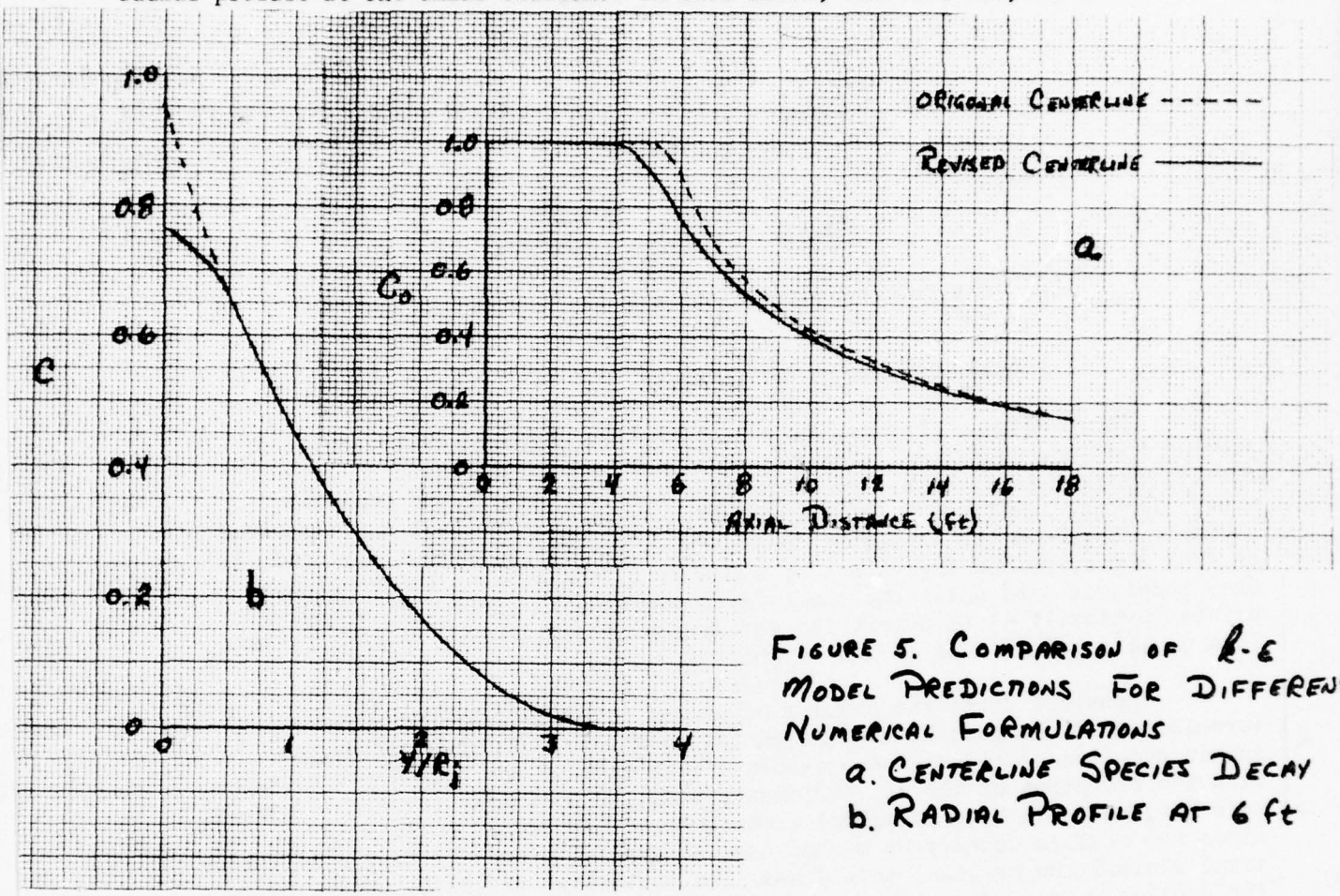


FIGURE 5. COMPARISON OF k-ε MODEL PREDICTIONS FOR DIFFERENT NUMERICAL FORMULATIONS
a. CENTERLINE SPECIES DECAY
b. RADIAL PROFILE AT 6 ft

point is tied to the local turbulence values at the point. The original formulation isolates the centerline diffusion from the rest of the flow field producing a larger core. The spiked concentration profile in the developing region also indicates this effect. The difference between the two predictions decreases with downstream distance.

While one can produce some logical reasons for believing in one centerline formulation rather than another, this is not always the case

when comparing various numerical techniques. Figure 6 illustrates the predictions of two computer codes using the one-equation

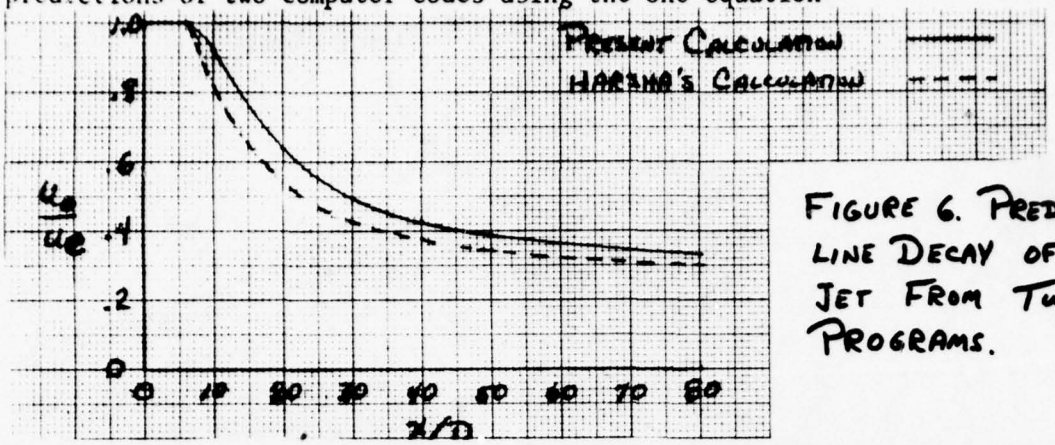


FIGURE 6. PREDICTED CENTER-LINE DECAY OF FORSTALL'S JET FROM TWO COMPUTER PROGRAMS.

TKE model. These are calculations of Forstall's experiment of an airjet into a coflowing air stream (NASA test case 9, Ref. 13). The differences between the two predictions in this case can be considered the numerical error inherent in the finite-difference precictions.

With this rough estimate of uncertainty in mind, we can proceed with a comparison of the prediction of the turbulence models with measurements of concentration and pitot pressure in a rocket exhaust. Figure 7a shows the predicted centerline decay of the concentration. All but the Ting-Libby correction appear to be a reasonable job in reproducing the data points. Given the experimental and numerical errors involved, one might be tempted to conclude that any of the mixing models, with the exception of Ting-Libby,[2] is adequate. If the infrared radiation from the plume were of ultimate interest, temperature is an important variable. The mixing models predict the centerline temperature distributions of Fig. 7b[1]. The difference in radiation computed from the three flow fields would be significant, much more so than a comparison such as Fig. 7a would indicate.

---

[1] The increase in temperature in the in region $0 \lesssim x \lesssim 2$ is probably a result of a readjustment of the exit plane values to equilibrium conditions.

[2] And the one-equation model as well.

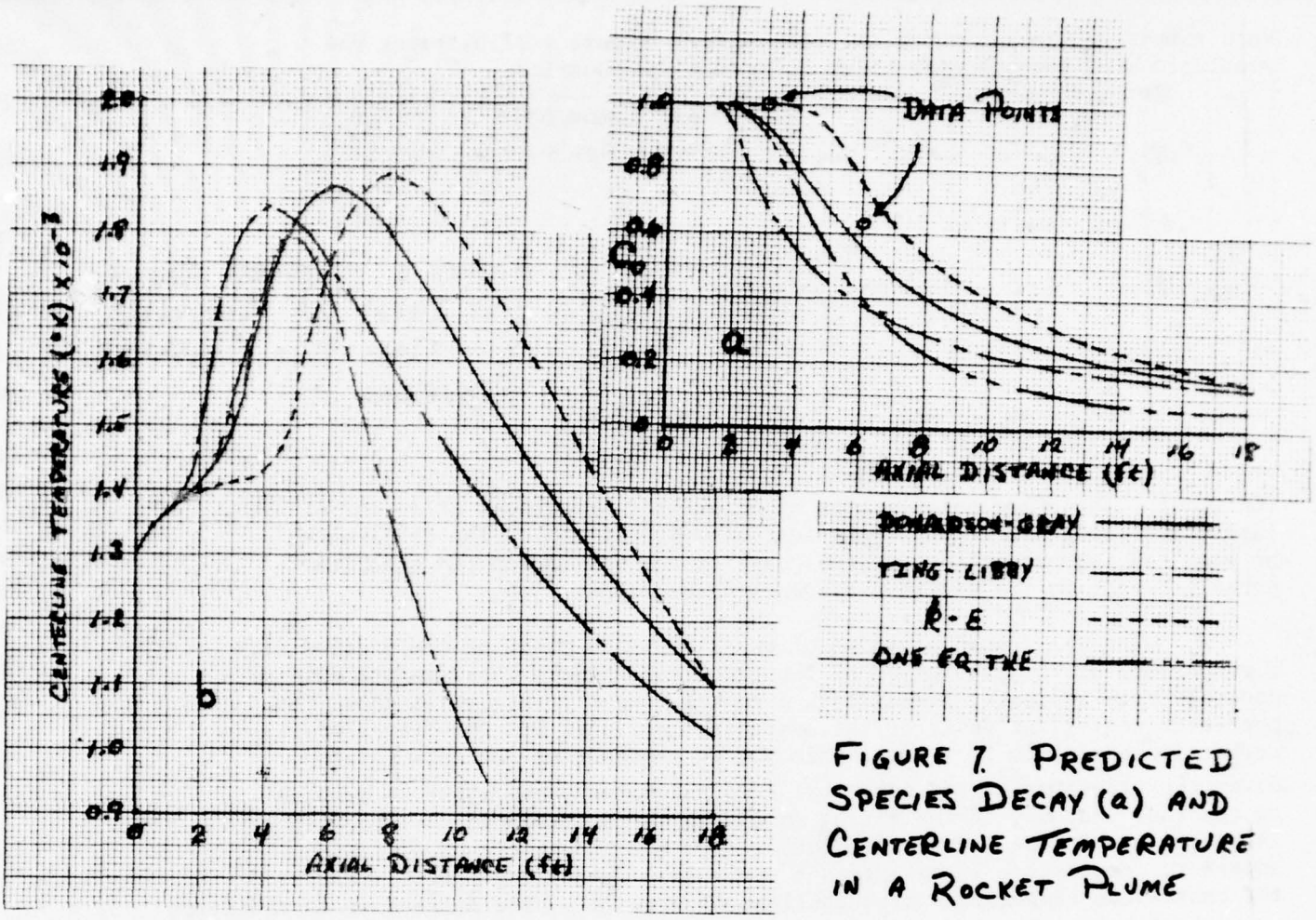


FIGURE 7. PREDICTED SPECIES DECAY (a) AND CENTERLINE TEMPERATURE IN A ROCKET PLUME

The radial profiles of concentration and pitot pressure at 3 and 6 ft from the exit plane are illustrated in Figs. 8 and 9. Pitot pressure was calculated from the Rayleigh supersonic pitot formula (Ref. 14). The comparisons do not reveal the clear-cut supremacy of any one of the mixing models. At 3 feet, the concentration profiles indicates that the mixing layer has not yet penetrated to the axis. The k-ε model is in very good agreement. However, the pitot pressure profile does not reveal a core region at this point. The Donaldson-Gray model looks to be the best. At the 6 ft station, none of the models represents the concentration profile, the Ting-Libby model while significantly underpredicting concentration is the best match to pitot pressure.

One of the interesting observations based on the comparisons of radial profiles is that the Prandtl model, equation (28), with either compressibility correction discussed above, gives profiles which are much more spread out than the to other models. There is more radial mixing with these models than with the other two as well as the data. This is a result of the minor variation of viscosity across the mixing region and the same viscosity in the freestream as in the plume. It might be expected that the prescription of a lower viscosity in the freestream than in the plume would tend to reduce the spread of radial profiles.

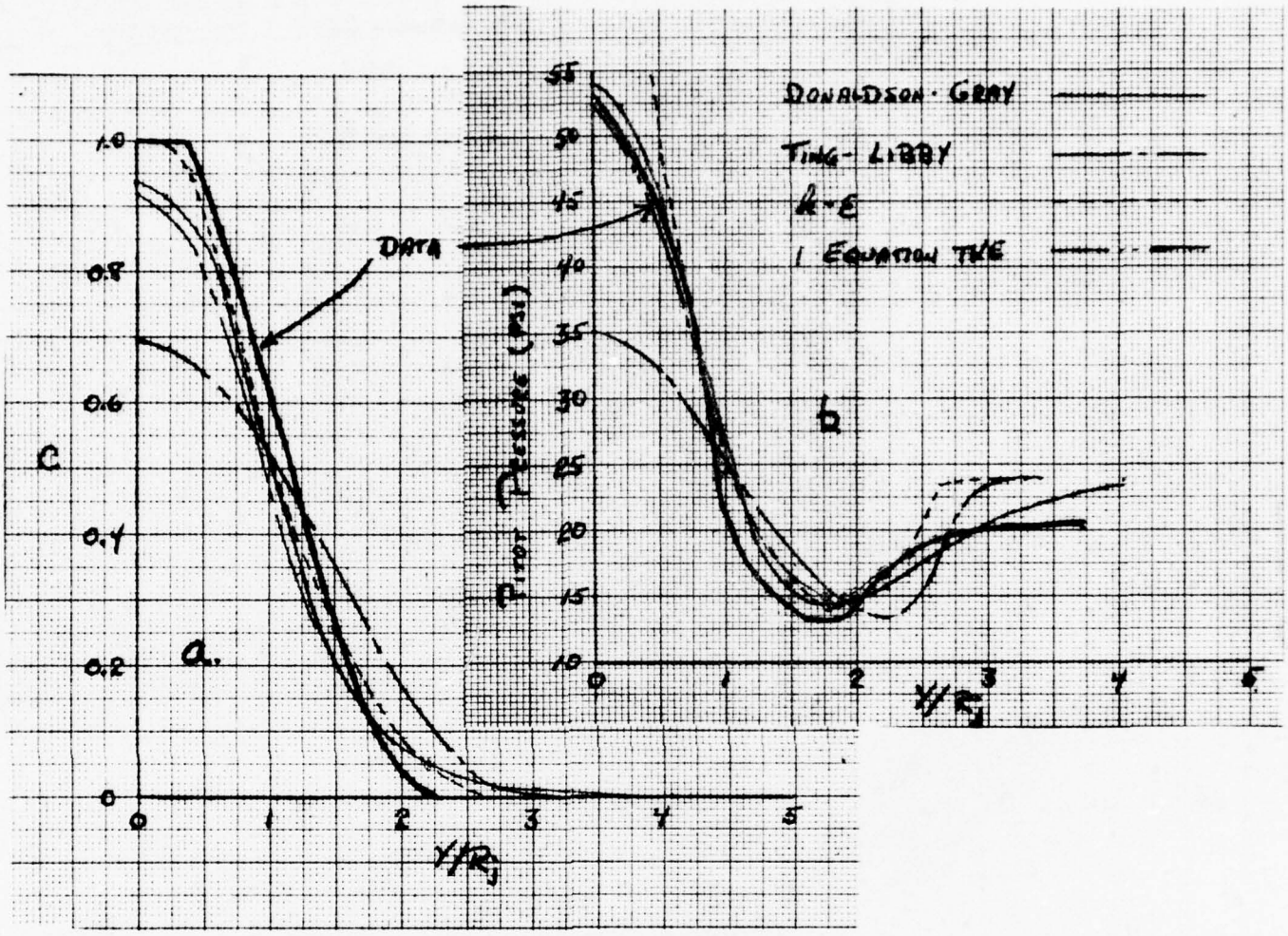


FIGURE 8. PREDICTED AND MEASURED RADIAL DISTRIBUTIONS OF SPECIES CONCENTRATION (a) AND PITOT PRESSURE (b) IN A ROCKET PLUME. $x = 3$ ft

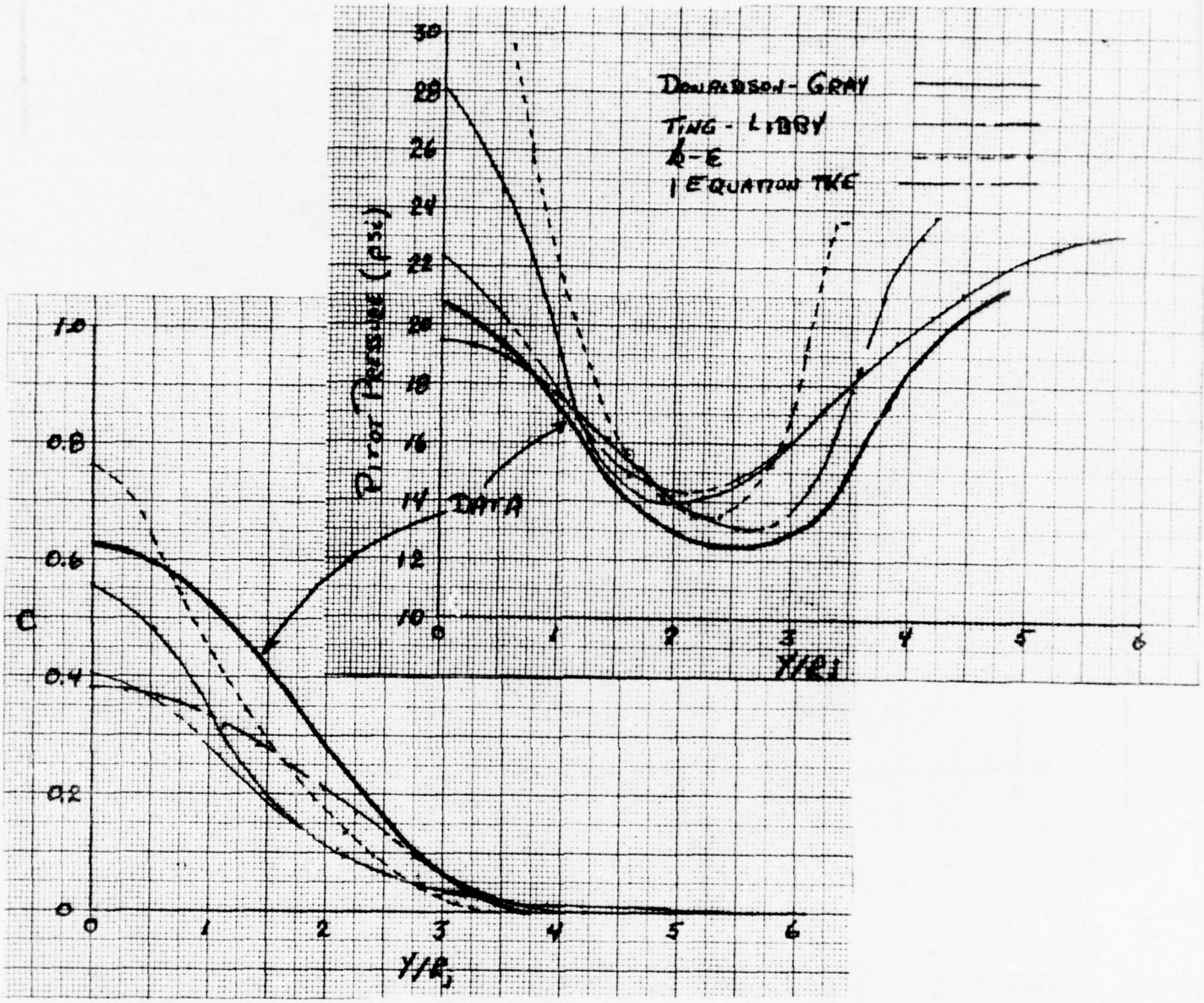


FIGURE 9. PREDICTED AND MEASURED RADIAL DISTRIBUTIONS OF SPECIES CONCENTRATION (a) AND PITOT PRESSURE (b) IN A ROCKET PLUME. $x = 6$ ft

A SIMPLIFIED LOW ALTITUDE PLUME MODEL

The two previous sections of this report have dealt with simplified idealizations of the low altitude plume. Emphasis was placed on the constant pressure mixing: how to arrive at a suitable description for the input conditions and how to model the mixing. This is the type of model most often used to calculate plumes. However, as mentioned previously, the initial portion of the plume is definitely not at constant pressure. The matched pressure point, once it is reached is not characterized by uniform temperature and velocity.

Recent attempts to formulate a more complete model for the plume have divided the flow field into an inviscid portion and a mixing layer which wraps around it. The two regions are assumed to be independent: there is no feedback of information from one region to the other. The inviscid flow is assumed to set the boundary conditions for the flow field. Such models involve several approximations. First, the pressure in the mixing region is assumed to be constant radially (with a value equal to the predicted interface pressure from the inviscid calculation). Second, there is in essence no penetration of the shock waves into the mixing region. Third, shock waves are not attenuated by the mixing region and progress downstream with very little damping. A priori judgement is required for when to end the inviscid calculation.

Unfortunately, there is no way to assess the magnitude of the errors caused by the assumptions in both of the above-mentioned plume models without solving the complete problem, which at present is apparently impossible. One might expect that the non-interaction of the shock waves with the mixing layer would lead to an overprediction of the plume temperature and hence other properties such as infrared radiation. On the other hand, the use of a uniform start-line in a constant pressure mixing program might underpredict the radiation since radiation is an exponential function of temperature.

Freeman (Ref. 15) suggested a procedure which could be used to assess these errors without having to solve the complete problem. He suggested an approximate solution to the exact problem, rather than an exact solution to an approximate problem. Such a model has definite merits. It is important to note that this type of model may result in poor predictions of the flow field and hence derived properties such as radiation.[1] The main utility of this model would lie in the determination of the relative magnitudes of the errors involved in the more exact solutions of the approximate models.

---

[1] Freeman's result (Ref. 15) indicate that the model may do a good job in this respect as well.

The model equations developed by Freeman, as will be shown below, are incorrect. In the remainder of this section we will develop the appropriate equations, indicate several different approximations which can be made and show how the equations can be solved.

Freeman's "Piecewise Uniform Flow Field" model envisions the plume as consiting of a number of discrete zones distributed radially. Each zone is characterized by uniform values of the dependent variables. The number of zones is, of course, arbitrary. Freeman proposed three. A one zone model would correspond to the start line procedure discussed above. As thenumber of zones increases, the model approaches an "exact" finite-difference form of the governing equations.

The three-zone model is shown in Fig. 10 . It would be a relatively simple procedure to extend the model to a more realistic four-zone model[1] in which the mixed region is devided into low speed (outer) and high speed (inner) zones. This, however,

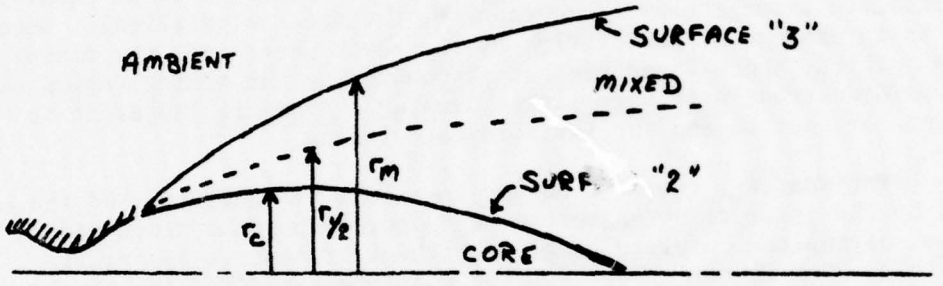


Figure 10. Definition of Piecewise Uniform Zones

detracts somewhat from the overall simplicity of the model by greatly increasing the algebra required to solve the resulting set of equations.

The development of the equations for each zone of the model starts with the general conservation equations (10) re-written here as:

$$\frac{\partial}{\partial x}(\rho u r) + \frac{\partial}{\partial r}(\rho v r) = 0$$

$$\frac{\partial}{\partial x}(\rho u^2 r) + \frac{\partial}{\partial r}(\rho u v r) = -\frac{\partial}{\partial x}(rp) + \frac{\partial}{\partial r}(r\tau_{xr})$$

$$\frac{\partial}{\partial x}(\rho u v r) + \frac{\partial}{\partial r}(\rho v^2 r) = -r\frac{\partial p}{\partial r} + \frac{\partial}{\partial r}(r\tau_{rr}) - \tau_{\theta\theta}$$

$$\frac{\partial}{\partial x}(\rho u H r) + \frac{\partial}{\partial r}(\rho v H r) = \frac{\partial}{\partial r}\left[\frac{\mu r}{Pr}\left(\frac{\partial H}{\partial r} + \frac{Pr-1}{2}\frac{\partial u^2}{\partial r}\right)\right] \qquad (31)$$

$$\frac{\partial}{\partial x}(\rho u \omega_x r) + \frac{\partial}{\partial r}(\rho v \omega_x r) = \frac{\partial}{\partial r}\left(\frac{\mu r}{Pr}\frac{\partial \omega_x}{\partial r}\right) + r R_\alpha$$

[1] The reason for making this extension will be obvious later.

In writing the equations in this fashion, we have made several assumptions. The right hand side of the energy equation is missing two terms (equivalent to Le = 1; σ_k = Pr). As shown below, these are of no consequence for the problem at hand. More important, we have assumed that transport in the radial direction is important, while transport in the axial direction is negligible compared to convection.

We now multiply each equation by dr and integrate between two arbitrary radial locations, r_1 and r_2. Liebnitz rule and some algebra give the result:

$$\frac{d}{dx}\int_{r_1}^{r_2} \rho u r\,dr - \rho u r\Big|_{r_2}\frac{dr_2}{dx} + \rho u r\Big|_{r_1}\frac{dr_1}{dx} = \rho v r\Big|_{r_1} - \rho v r\Big|_{r_2}$$

$$\frac{d}{dx}\int_{r_1}^{r_2} \rho u^2 r\,dr - u_2\frac{d}{dx}\int_0^{r_2}\rho u r\,dr + u_1\frac{d}{dx}\int_0^{r_1}\rho u r\,dr = -\frac{d}{dx}\int_{r_1}^{r_2} P r\,dr + P r\Big|_{r_2}\frac{dr_2}{dx} - P r\Big|_{r_1}\frac{dr_1}{dx} + r\tau_{xr}\Big|_{r_2} - r\tau_{xr}\Big|_{r_1}$$

$$\frac{d}{dx}\int_{r_1}^{r_2} \rho u r v\,dr - v_2\frac{d}{dx}\int_0^{r_2}\rho u r\,dr + v_1\frac{d}{dx}\int_0^{r_1}\rho u r\,dr = -rP\Big|_{r_2} + rP\Big|_{r_1} + \int_{r_1}^{r_2} P\,dr + r\tau_{rr}\Big|_{r_2} - r\tau_{rr}\Big|_{r_1} - \int_{r_1}^{r_2}\tau_{\theta\theta}\,dr \qquad (32)$$

$$\frac{d}{dx}\int_{r_1}^{r_2} \rho u H r\,dr - H_2\frac{d}{dx}\int_0^{r_2}\rho u r\,dr + H_1\frac{d}{dx}\int_0^{r_1}\rho u r\,dr = \frac{\mu r}{P_r}\frac{\partial H}{\partial r}\Big|_{r_1}^{r_2} + \frac{\mu r}{P_r}\frac{P_r - 1}{2}\frac{\partial u^2}{\partial r}\Big|_{r_1}^{r_2}$$

$$\frac{d}{dx}\int_{r_1}^{r_2} \rho u \omega_x r\,dr - \omega_{x2}\frac{d}{dx}\int_0^{r_2}\rho u r\,dr + \omega_{x1}\int_0^{r_1}\rho u r\,dr = \frac{\mu r}{P_r}\frac{\partial \omega_x}{\partial r}\Big|_{r_1}^{r_2} + \int_{r_1}^{r_2} R_x r\,dr$$

Equations (32) are generalized and simplified by applying them to each of the three zones of Fig. 10 . We assume that each zone is characterized by some uniform, average value of the dependent variables (ρ, u, v, P, etc.) at each axial location. Hence the integrals appearing in the equations can be simply expressed. We also note that at the boundary surfaces of the mixing layer (r_c, r_m) the turbulent viscosity, and hence the shear stress, is zero since by definition this is the place where there is no turbulence. We assume the mean value of $\tau_{\theta\theta}$ in the mixing region is zero, since it is probably negligibly small compared to the radial pressure.

Each of the five conservation equations can be applied to each of the three zones, as well to all three simultaneously. This gives twenty equations, only fifteen of which are independent. They are:

MASS

Core: $$\frac{d\dot{m}_c}{dx} = 2\pi\rho_2 r_c\left(u_2\frac{dr_c}{dx} - v_2\right)$$

Mixed: $$\frac{d\dot{m}_m}{dx} = 2\pi\left[\rho_2 r_c\left(v_2 - u_2\frac{dr_c}{dx}\right) + \rho_3 r_m\left(u_3\frac{dr_m}{dx} - v_3\right)\right]$$

Ambient: $$\frac{d\dot{m}_a}{dx} = 2\pi\rho_3 r_m\left(v_3 - u_3\frac{dr_m}{dx}\right)$$

Total: $$\frac{d}{dx}(\dot{m}_c + \dot{m}_m + \dot{m}_a) = 0 \qquad (33)$$

AXIAL MOMENTUM

Core: $$\frac{d}{dx}\left(\dot{m}_c u_c + \pi P_c r_c^2\right) = u_2\frac{d\dot{m}_c}{dx} + 2\pi P_2 r_c\frac{dr_c}{dx}$$

Mixed: $$\frac{d}{dx}\left[\dot{m}_m u_m + \pi P_m\left(r_m^2 - r_c^2\right)\right] = u_3\frac{d}{dx}(\dot{m}_c + \dot{m}_m) - u_2\frac{d\dot{m}_c}{dx} + 2\pi\left[P_3 r_m\frac{dr_m}{dx} - P_2 r_c\frac{dr_c}{dx}\right] \qquad (34)$$

Ambient: $$\frac{d}{dx}\left[\dot{m}_a u_a + \pi P_a\left(r_a^2 - r_m^2\right)\right] = -u_3\frac{d}{dx}(\dot{m}_c + \dot{m}_m) - 2\pi P_3 r_m\frac{dr_m}{dx}$$

Total: $$\frac{d}{dx}\left[\dot{m}_c u_c + \dot{m}_m u_m + \dot{m}_a u_a + \pi\left(P_c r_c^2 + P_m\left(r_m^2 - r_c^2\right) + P_a\left(r_a^2 - r_m^2\right)\right)\right] = 0$$

RADIAL MOMENTUM

Core: $$\frac{d}{dx}\dot{m}_c v_c = v_2\frac{d\dot{m}_c}{dx} + 2\pi r_c\left(P_c - P_2\right)$$

Mixed: $$\frac{d}{dx}\dot{m}_m v_m = v_3\frac{d}{dx}(\dot{m}_c + \dot{m}_m) - v_2\frac{d\dot{m}_c}{dx} + 2\pi\left[r_m\left(P_m - P_3\right) + r_c\left(P_2 - P_m\right)\right]$$

Ambient: $$\frac{d}{dx}\dot{m}_a v_a = -v_3\frac{d}{dx}(\dot{m}_c + \dot{m}_m) + 2\pi r_m\left(P_3 - P_a\right) \qquad (35)$$

Total: $$\frac{d}{dx}\left[v_c\dot{m}_c + v_m\dot{m}_m + v_a\dot{m}_a\right] = 2\pi\left[r_2\left(P_c - P_m\right) + r_m\left(P_m - P_a\right)\right]$$

ENERGY

$$\text{Core:} \quad \frac{d}{dx}(\dot{m}_c H_c) - H_2 \frac{d\dot{m}_c}{dx} = 0$$

$$\text{Mixed:} \quad \frac{d}{dx}(\dot{m}_m H_m) - H_3 \frac{d}{dx}(\dot{m}_c + \dot{m}_m) + H_2 \frac{d\dot{m}_m}{dx} = 0$$

$$\text{Ambient:} \quad \frac{d}{dx}(\dot{m}_a H_a) + H_3 \frac{d}{dx}(\dot{m}_c + \dot{m}_m) = 0$$

$$\text{Total:} \quad \frac{d}{dx}\left[\dot{m}_c H_c + \dot{m}_m H_m + \dot{m}_a H_a\right] = 0 \qquad (36)$$

SPECIES

$$\text{Core:} \quad \frac{d}{dx}(\dot{m}_c \omega_{kc}) - \omega_{k2} \frac{d\dot{m}_c}{dx} = \pi R_{kc} r_c^2$$

$$\text{Mixed:} \quad \frac{d}{dx}(\dot{m}_m \omega_{km}) - \omega_{k3} \frac{d}{dx}(\dot{m}_c + \dot{m}_m) + \omega_{k2} \frac{d\dot{m}_c}{dx} = \pi R_{km} (r_m^2 - r_c^2)$$

$$\text{Ambient:} \quad \frac{d}{dx}(\dot{m}_a \omega_{ka}) + \omega_{k3} \frac{d}{dx}(\dot{m}_c + \dot{m}_m) = \pi R_{ka} (r_a^2 - r_m^2) \qquad (37)$$

$$\text{Total:} \quad \frac{d}{dx}\left[\dot{m}_c \omega_{kc} + \dot{m}_m \omega_{km} + \dot{m}_a \omega_{ka}\right] = \pi\left[r_c^2 (R_{kc} - R_{km}) + r_m^2 (R_{km} - R_{ka}) + r_a^2 R_{ka}\right]$$

The mass flow rates are given by:

$$\begin{aligned} \dot{m}_c &= \pi \rho_c u_c r_c^2 \\ \dot{m}_m &= \pi \rho_m u_m (r_m^2 - r_c^2) \\ \dot{m}_a &= \pi \rho_a u_a (r_a^2 - r_m^2) \end{aligned} \qquad (38)$$

Equation $(38)_3$ may be thought of as defining r_a in terms of $\dot{m}_a$, which is more convenient for missiles in flight.

In addition, we have the ideal gas relations:

$$\rho_i = \frac{P_i}{R_i T_i} \qquad (i = c, m, a) \qquad (39)$$

and relationships between total enthalpy and temperature:

$$
\begin{aligned}
H_i &= h_i + \tfrac{1}{2}(u_i^2 + v_i^2) \\
&= \sum \left[\omega_{\alpha i} h^{\circ}_{\alpha i} + \int_{T_0}^{T_i} c_{p\alpha i}\, \omega_{\alpha i}\, dT \right] + \tfrac{1}{2}(u_i^2 + v_i^2) \\
&= \sum (\omega_{\alpha i} h^{\circ}_{\alpha i}) + \int_{T_0}^{T_i} c_{p_i}\, dT + \tfrac{1}{2}(u_i^2 + v_i^2) \\
&= \sum \omega_{\alpha i} h^{\circ}_{\alpha i} + \bar{c}_{p i}(T_i - T_0) + \tfrac{1}{2}(u_i^2 + v_i^2)
\end{aligned}
\qquad (40)
$$

$$(i = c, m, a)$$

Despite the fact that we have available some 24 equations, we are still in need of further information: there are 39 unknowns , assuming average reaction rates and species molecular weights, heats of formation and heat capacities are known.

Before discussing how the set of equations can be closed, it is worthwhile to point out how this set differs from the equations of Ref. 15 . There are more equations provided in the present development; although neither set of equations is closed at the present time. Freeman uses shear stress defined at the boundaries of the mixing layer. As pointed out above the stress at these locations is zero. Freeman's "stresses" are actually entrainment coefficients, empirical models for the growth of the mixing region by capturing mass from the freestream and the core. Finally, the radial momentum equations differ since the equations of Ref. 15 ignore the increase in radial momentum due to the increase in size of the layer.

Returning now to the problem at hand, we see that equations (35) through (37) must be supplemented by additional relations. Our unknowns contain values at the boundaries of the layer in addition to those within the layer. We anticipate that these should be related. For example, the value in the region might be taken as the average of the properties at the boundary, or be value at the boundary might be defined in terms of the values on either side of the boundary. Obviously several alternatives are possible. We realize that there will be large pressure and radial velocity gradients across the zones. We will assume we know these values at the boundaries and define:

$$
\begin{aligned}
v_m &\equiv \frac{v_2 + v_3}{2} \\
P_m &\equiv \frac{P_2 + P_3}{2}
\end{aligned}
\qquad (41)
$$

Defined in this manner, we expect $v_c = o = v_a$. Also, we would expect the axial velocity to be somewhat more uniform across each zone. Therefore, we take[1]:

$$u_2 \equiv \frac{u_c + u_m}{2}$$
$$u_3 \equiv \frac{u_a + u_m}{2}. \tag{42}$$

We assume that the freestream values are known and uniform:

$$\begin{aligned} u_a &= \text{constant} \\ H_a &= \text{constant} \\ \omega_{\alpha a} &= \text{constant} \\ P_a &= \text{constant} \\ \rho_a &= \text{constant} \end{aligned} \tag{43}$$

and that the core total enthalpy is constant:

$$\begin{aligned} \omega_{\alpha c} &= \text{constant} \\ H_c &= \text{constant}. \end{aligned} \tag{44}$$

With H_a and H_c constant, we find from $(36)_1$ and $(36)_3$

$$H_2 = H_c;\ H_3 = H_a \tag{45}$$

At the plume freestream boundary, we will assume the pressure to be a constant for subsonic external flow or to be given by the Newtonian impact pressure for supersonic external flow:

$$P_3 = \begin{cases} P_a & M_\infty < 1 \\ P_a\left(1 + \gamma_\infty M_\infty^2 \sin^2 \tan^{-1}\frac{dr_m}{dx}\right) & \\ \equiv P_a\left(1 + \gamma_\infty M_\infty^2 \sin^2 \frac{dr_m}{dx}\right) & M_\infty > 1 \end{cases} \tag{46}$$

---

[1] An alternate procedure for the boundary velocities would be to use Bernoulli's equation or an equivalent to calculate u_{r_c} and u_{r_m} knowing the pressure history on the surface.

This increases the number of equations to 37. To obtain the final two equaties we need two additional assumptions. First we assume some prescribed total pressure distribution in the core flow:

$$\rho_c u_c^2 + P_c = P_t(x) \equiv f(x)$$

$$\rho_c u_c \frac{du_c}{dx} + \frac{dP_c}{dx} = f'(x) \qquad (47)$$

Equation $(47)_2$ is in essence a mechanical energy balance or Bernoulli equation for the core. The variation of total pressure with axial distance can be taken as a continuous one, where the shock structure is smeared over the entire core; zero, for isentropic or shock-free core flow; or as a discrete function of axial distance, where some "mass-averaged" shock is assumed to occur at a position where the core radius, for example, is a local maximum.

The final equation is obtained by introducing the concept of a shear stress. There certainly is a stress inside the mixed region, so we write an axial momentum balance from the centerline to the mixed region half-radius, $r_{1/2}$:

$$\frac{d}{dx}(\dot{m}_c u_c + \dot{m}_{1/2} u_{1/2}) - u_{1/2} \frac{d}{dx}(\dot{m}_c + \dot{m}_{1/2}) = -\pi \frac{d}{dx}\left[P_c r_c^2 + P_{1/2}(r_{1/2}^2 - r_c^2)\right]$$

$$+ 2\pi\left[P_{1/2} r_{1/2} \frac{dr_{1/2}}{dx} + r_{1/2} \tau_{xr}\right]. \qquad (48)$$

If we assume the half radius is exactly half way into the mixed region:

$$r_{1/2} \equiv \frac{r_c + r_m}{2} \qquad P_{1/2} \equiv P_m$$

$$\dot{m}_{1/2} \equiv \frac{\dot{m}_m}{2} \qquad u_{1/2} \equiv u_m \; . \qquad (49)$$

The shear stress can be taken as:

$$\tau_{xr} \equiv \mu_t \frac{du_m}{dr}$$

$$\equiv \mu_t \frac{u_{r_m} - u_{r_c}}{r_m - r_c}$$

$$= \rho_{1/2}\, c (r_m - r_c) \left|u_{r_m} - u_{r_c}\right| \frac{(u_{r_m} - u_{r_c})}{r_m - r_c}$$

$$= -\rho_{1/2}\, c\, (u_{r_c} - u_{r_m})^2 . \qquad (50)$$

We now have a determinate problem. Equations (33) through (37) are re-written as:

$$\frac{d\dot{m}_c}{dx} = 2\pi \rho_2 r_c \left(u_2 \frac{dr_c}{dx} - v_2\right)$$

$$\frac{d\dot{m}_a}{dx} = 2\pi \rho_3 r_m \left(v_3 - u_3 \frac{dr_m}{dx}\right)$$

$$\frac{d}{dx}\left(\dot{m}_c + \dot{m}_m + \dot{m}_a\right) = 0$$

$$\frac{d}{dx}\left(\dot{m}_c u_c + \pi P_c r_c^2\right) = u_2 \frac{d\dot{m}_c}{dx} + 2\pi r_c P_2 \frac{dr_c}{dx}$$

$$\left(u_a - u_3 + \frac{P_a}{\rho_a u_a}\right) \frac{d\dot{m}_a}{dx} = -2\pi P_3 r_m \frac{dr_m}{dx}$$

$$\frac{d}{dx}\left[\dot{m}_c u_c + \dot{m}_m u_m + \dot{m}_a\left(u_a + \frac{P_a}{\rho_a u_a}\right) + \pi r_c^2\left(P_c - \frac{P_2+P_3}{2}\right) + \pi r_m^2 \frac{P_2+P_3}{2}\right] = 0$$

$$v_2 \frac{d\dot{m}_c}{dx} = -2\pi r_c \left(P_c - P_2\right)$$

$$v_3 \frac{d\dot{m}_a}{dx} = 2\pi r_m \left(P_a - P_3\right) \qquad (51)$$

$$\frac{d}{dx}\left[\dot{m}_c H_c + \dot{m}_m H_m + \dot{m}_a H_a\right] = 0$$

$$\frac{d}{dx}\left(\omega_{xm} \dot{m}_m\right) + \omega_{xa} \frac{d\dot{m}_a}{dx} + \omega_{xc} \frac{d\dot{m}_c}{dx} = \pi R_{xm} \left(r_m^2 - r_c^2\right)$$

These ten coupled equations are the one which must be solved.

The solution procedure can be simplified with some algebraic manipulations. It is straight forward, though tedious, to reduce equations (51) to the following set:

$$\frac{dr_3}{dx} = \frac{P_3 - P_a + \rho_3 v_3^2}{\rho_3 v_3 u_3}$$

$$P_3 = \begin{cases} P_a & M_\infty < 1 \\ P_a\left(1 + \gamma_\infty M_\infty^2 \sin^2 \frac{dr_m}{dx}\right) & M_\infty > 1 \end{cases}$$

$$u_3 = u_a \qquad M_\infty < 1$$

$$\dot{m}_m + \dot{m}_c + \dot{m}_a = C_1$$

$$\dot{m}_c u_c + \dot{m}_a\left(u_a + \frac{P_a}{\rho_a u_a}\right) + \dot{m}_m u_m + \pi r_c^2\left(P_c - \frac{P_2+P_3}{2}\right) + \pi r_m^2 \frac{P_2+P_3}{2} = C_2$$

$$\frac{dr_2}{dx} = \frac{P_3 - P_2 + \rho_2 v_2^2}{\rho_2 v_2 u_2}$$

$$v_3 = \pm\left\{\left[\left(u_a - u_3 + \frac{P_a}{\rho_a u_a}\right)\frac{\rho_3 u_3}{P_3} - 1\right]\frac{P_3 - P_a}{\rho_3}\right\}^{1/2}$$

$$\frac{d\dot{m}_c}{dx} = 2\pi r_2\left[\frac{P_2 - P_c}{v_2}\right]$$

$$\frac{d\dot{m}_a}{dx} = \begin{cases} -2\pi\rho_a r_m u_a \dfrac{dr_m}{dx} & M_\infty < 1 \\ 2\pi r_3 \dfrac{P_a - P_3}{v_3} & M_\infty > 1 \end{cases} \qquad (52)$$

$$u_c = u_2 + \frac{P_2 - P_c + \rho_2 v_2^2}{\rho_2 u_2} - \frac{f'(x)\, r_c}{2(P_2 - P_c)}$$

$$\frac{\dot{m}_m}{2}\frac{du_m}{dx} + (u_c - u_m)\,2\pi r_c \frac{P_2 - P_c}{v_2} + \pi r_c^2 f'(x) - 2\pi r_c (P_m - P_c)\frac{P_2 - P_c + \rho_2 v_2^2}{\rho_2 v_2 u_2} + \pi\left(r_{1/2}^2 - r_2^2\right)\frac{dP_m}{dx} - 2\pi r_{1/2}\tau_{xr} = 0$$

$$\rho_c u_c^2 + P_c = f(x)$$

For supersonic external flows, we can look on $(52)_4$ as providing an expression for $\frac{du_m}{dx}$. Then $(52)_{11}$ contains only a single derivative. For subsonic external flows, $v_3 = 0$ and equation $(52)_1$ is indeterminate. In this case, we should take

$$u_m = \frac{u_a + u_2}{2}, \qquad (53)$$

find $\dot{m}_m$ from $(52)_2$ and in a from $(52)_1$, again giving a set of equations which contain only one derivative in each.

The resulting set of coupled, non-linear equations must be solved numerically. There are, however, a variety of techniques for obtaining a solution (eg. Ref. 3). Unfortunately, there was insufficient time to set up and solve the numerical equations.

CONCLUSIONS AND RECOMMENDATIONS

Most of the major conclusions and recommendations from this investigation are presented in the appropriate sections. They are summerized below.

SIMPLE EXPANSION MODEL

1. The differences between perfect gas and real gas start-line properties is small. The perfect gas calculation is significantly simpler to perform. The resulting errors are probably of minor importance compared to the errors involved in the lumping procedure.

2. The axial location to which the predicted start-line properties correspond is impossible to determine. When an overall property, such as total radiation, is the ultimate goal, this indeterminacy is unimportant. When detailed spatial data is in interest, the actual exit plane is probably the appropriate place.

3. Since the actual "constant pressure line" in the plume is characterized by a distribution of temperatures rather than a single value, the effect of mixing on the downstream properties must be investigated. There is conflicting evidence as to the importance of a variable start-line on downstream properties.

EVALUATION OF TURBULENCE MODELS

1. In a single computer code, slight variations in numerical formulations usually result in minor changes in the overall predictions. However, the variation from one code to another can be significant. These variations should be examined in more detail in order to appreciate the magnitude of the numerical uncertainties in resulting answers.

2. None of the turbulence models investigated here provides a good description of turbulent mixing in rocket exhausts. Further investigations into this problem are required, examining properties like emitted radiation in addition to species concentrations and pitot pressure.

3. The data base on turbulent mixing in a supersonic jet in a supersonic freestream should be expanded to allow for additional comparisons.

4. Compressibility effects in the turbulence models should be examined in more detail.

SIMPLIFIED LOW ALTITUDE MODEL

1. The governing equations should be solved numerically and a parametric study undertaken to determine the relative importance of shocks, lateral pressure gradients and mixing on the overall properties of the plume. Such a study may help to identify critical areas of the fluid mechanics which should be modeled to detail as well as less important areas which can be handled approximately. Failures in current models will not necessarily be removed by the addition of more and more complex components. In a problem as complex as the modeling of a rocket plume there must of necessity be certain components omitted. The continual "improvement" of plume models may make them more complex without provided better predictions

REFERENCES

1. Sukanek, P. C., "A Simple Method of Calculating the Start-Line Properties of Constant Pressure Rocket Exhausts," Submitted for Publication, AIAAJ .

2. Jarvenin, P. O. and J. A. F. Hill, "Universal Model for Underexpanded Rocket Plumes in Hypersonic Flow," 12th JANNAF Liquid Propulsion Meetin, Las Vegas (1970).

3. Young, D. M. and R. T. Gregory, A Survey of Numerical Mathematics, Vol. I, Addison-Wesley, Reading (1972).

4. Mikatarian, R. R., C. J. Kau and H. S. Pergament, "A Fast Computer Program for Nonequilibrium Rocket Plume Predictions, " AFRPL-TR-72-94, (1972).

5. Patankar, S. V. and D. B. Spalding, Heat and Mass Transfer in Boundary Layers, Intertext, London (1970).

6. Prandtl, S. "Bremerkungen zu Theori der frein Turbulenz," ZAMM 22, (1942).

7. Harsha, P. T., "A General Analysis of Free Turbulent Mixing," AEDC-TR-73-177, (1974).

8. Launder, B. E., A. Morse, W. Rodi, and D. B. Spalding, "Prediction of Free Shear Flows. A Comparison of the Performance of Six Turbulence Models," in Free Turbulent Flows, Vol. I, NASA SP-321 (1973).

9. Donaldson, C. dux P. and K. E. Gray, "Theoretical and Experimental Study of the Compressible Free Mixing of Two Dissimilar Gases," AIAAJ 4, 2017 (1966).

10. Smoot, L. D., "Turbulent Mixing Coefficients for Compressible Coaxial Submerged and Coflowing Jets," AIAAJ 14, 1699 (1976).

11. L. Ting and P. A. Libby, "Remarks on the Eddy Viscosity in Compressible Mixing Flows," J. Aero. Sci. 27, 797 (1960).

12. Dash, S. G. Weilerstein and R. Vaglio-Lauren, "Compressibility Effects in Free Turbulent Shear Flows," AFOSR-TR-75-1436 (1975).

13. "Free Turbulent Shear Flows," Vol. II, Summary of Data.

14. Liepmann, H. W. and A. Roshko, Elements of Gas Dynamics , Wiley, New York, p. 149 (1957).

15. Freeman, G. N., "An Alternate Model for Low Altitude Plumes," Paper presented at the JANNAF Low Altitude Plume Technology Workshop, Redstone Arsenal (1977).

1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM

sponsored by

THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

conducted by

AUBURN UNIVERSITY AND THE OHIO STATE UNIVERSITY

PARTICIPANT'S FINAL REPORT

SIMULATION OF TURBULENT FLOW VELOCITY

Prepared by:	John W. Weatherly, III, Ph.D
Academic Rank:	Associate Professor
Department and University:	Department of Mechanical Engineering Louisiana State University Baton Rouge, Louisiana
Assignment:	Arnold Engineering Development Center Engine Test Facility Arnold Air Force Station, TN
Research Colleagues:	D. O. Barnett and T. V. Giel
Date:	August 12, 1977
Contract No.:	F44620-75-C-0031

SIMULATION OF TURBULENT FLOW VELOCITY

ABSTRACT

Laser velocimeters can be used to measure fluid velocities by measuring particle velocities in the fluid flow field. If the particles do not follow the flow then the laser velocimeter measurements will not accurately represent the fluid velocity. Particle dynamic calculations can be used to define this inaccuracy if the space-time history of the flow field is known. This study investigated methods of simulating the space-time history of a turbulent flow from the gross statistics of the velocity so that the inaccuracy of laser velocimeter turbulent flow measurements might be estimated.

ACKNOWLEDGMENT

The author is sincerely grateful to the Air Force Systems Command for support of this summer research. A deep appreciation is due to ASEE and Auburn University and in particular Mr. Fred O'Brien and Dr. M. L. Laster for excellent administration of the project. Their work has certainly made this summer more productive and enjoyable.

The author is indebted to the Arnold Engineering Development Center for providing a cordial and helpful work environment. It has been a pleasure and a professionally gratifying experience to work with these very capable people. In particular, thanks are due Mr. Marshall K. Kingery and Mrs. Kathy White.

Also, it has been enjoyable and profitable working with the staff of ARO, Inc, namely Dr. D. O. Barnett, Mr. T. V. Giel and Mr. F. C. Loper. I would also like to express my thanks to Dr. W. Frost, Dr. G. H. Fichtl, Dr. S. I. Cheng, Dr. G. D. Whitehouse, Professor A. J. McPhate and Mr. Claude Mount for their help with this problem.

Introduction

Laser velocimeters (LV) measure fluid velocities by measuring the velocities of particles suspended in the fluid. If the particles follow the fluid flow then an LV will accurately measure this flow. In the ideal case of small particles in a flow having only slow velocity changes this situation is closely approached. To the contrary, in flows where the velocity changes rapidly and large heavy particles are being measured, a considerable difference between the fluid motion and the particle motion will result.

To assess the accuracy of LV measurements the difference between the fluid motion and particle motions must be determined for the various types of particles which can be detected by the LV. The difference between the fluid and particle motions can be predicted using drag relations, such as Stoke's drag, if the fluid velocity space-time history is known. Unfortunately the velocity history is not known for turbulent flows which fluctuate rapidly. The accuracy of LV turbulent flow measurements might be estimated, however, if the fluid velocity history can be simulated with some pseudo-random function which has the correct (1) mean velocity, (2) root mean square velocity (r.m.s), and (3) auto-correlation function. The purpose of this study was to pick such a pseudo-random function which can be given a predetermined mean, r.m.s and auto-correlation as a function of location in the flow.

Laser Velocimetry

A one-component fringe system with crossed beam optics is shown below:

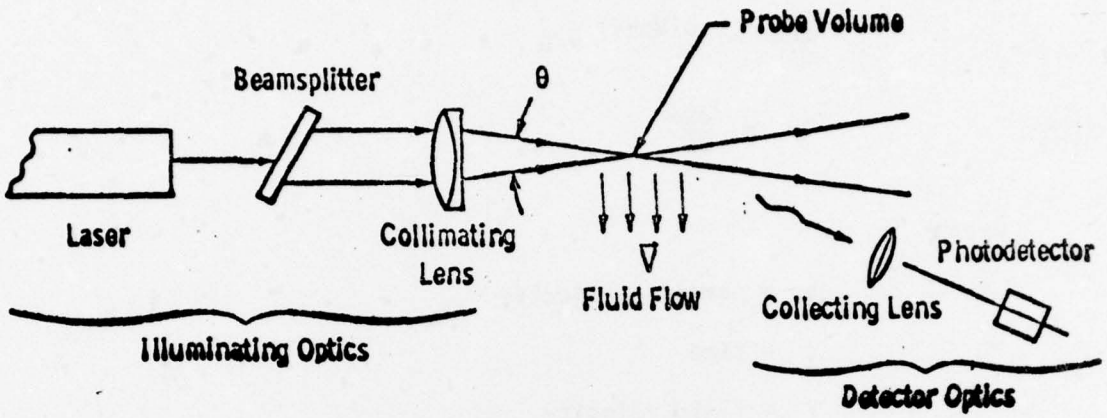


Figure 1. Dual Scatter Laser Velocimeter

Light from a laser is split into two beams and focused at a point in the flow field.[1] With this arrangement, interference fringes are created at the probe volume in the flow field. Particles in the fluid which are carried through the fringes scatter light which is collected by the detector optics and focused on a photo detector. The signal produced by the photo detector is analyzed to give the frequency at which a particle intercepts the fringes. The velocity of each particle is related to the signal frequency (f_i) by the relation

$$V_i = Kf_i \tag{1}$$

where K is the fringe spacing. To determine the fluid velocity, a relationship between the particle velocity and the fluid velocity must be given.

Particle Dynamics & Fluid Flow

The equations of motion for a particle suspended in a fluid are:[2]

$$\frac{dVp}{dt} + \sigma(Vp-V) = 0 \tag{2}$$

$$Vp = \frac{dXp}{dt} \tag{3}$$

where

Vp = particle velocity

t = time

V = fluid velocity

Xp = particle position

The response parameter, σ, is defined as:

$$\sigma = \frac{18\mu}{\rho_p D_p^2} \tag{4}$$

where

μ = fluid viscosity

ρ_p = particle density

D_p = particle diameter

Figure 2[3] shows the variations of σ with D_p and ρ_p. For a large σ (i.e. small diameter or small density or large fluid viscosity) the particle should follow the fluid flow reasonably well. For a small σ the reverse will be true.

If the fluid velocity is known as a function of time and space, equations (2) and (3) can be solved. For example, in a fluid flowing through a normal shock particle velocities can be calculated yielding a response such as is shown in Figure 3.[3]

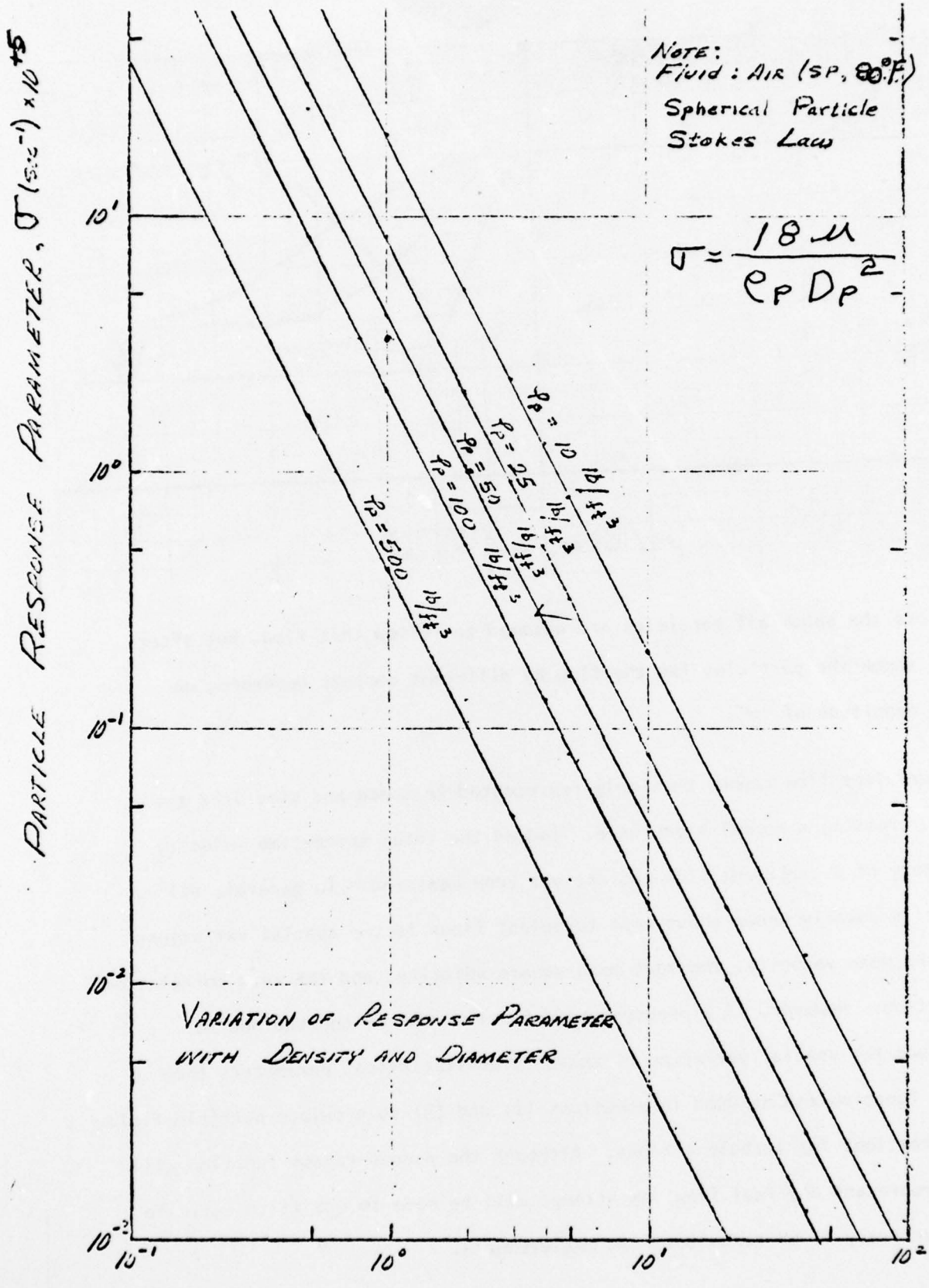


FIG. 2 PARTICLE DIAMETER, d_p (μm)

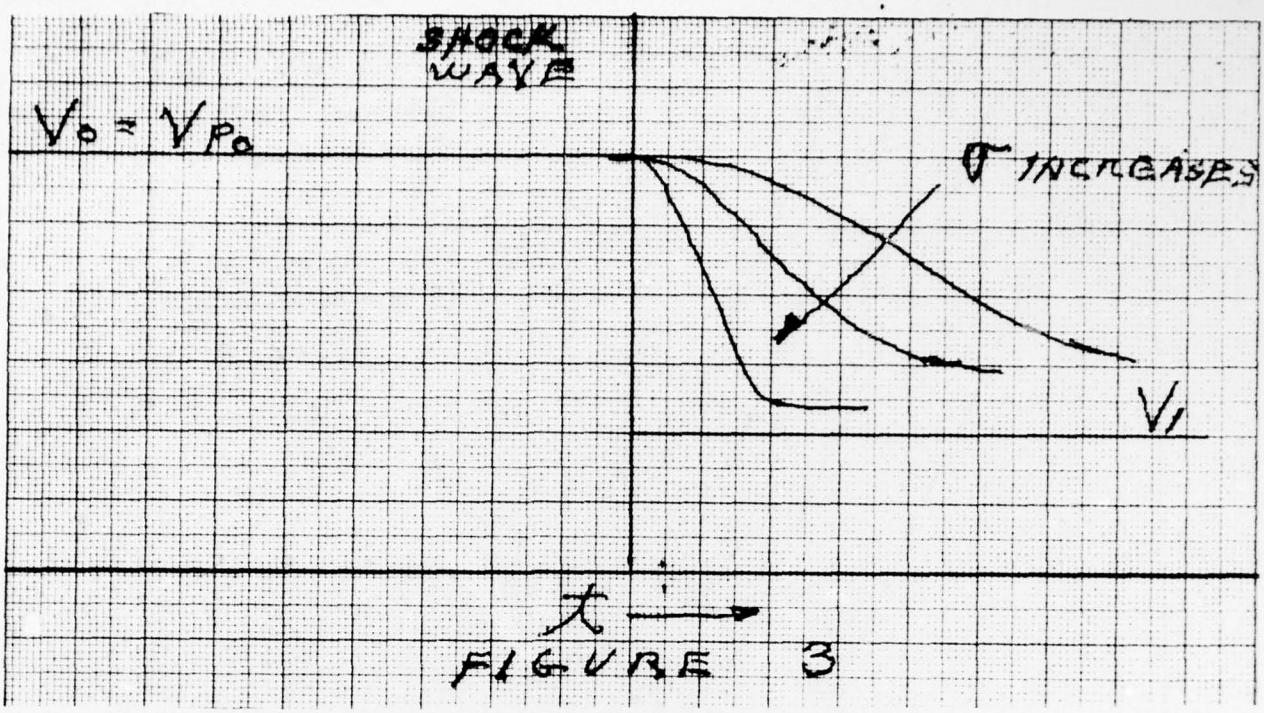


FIGURE 3

Before the shock all particles are assumed to follow this flow, but after the shock the particles lag the flow by different amounts depending on the magnitude of "σ".

A turbulent flow cannot be simply represented in space and time like the flow crossing a normal shock wave. Indeed the total space-time velocity history of a turbulent flow has not yet been measured. In general, all that is usually known about most turbulent flows is the spatial variations of the mean velocity, the root mean square velocity, and the auto correlation function. However, if a pseudo-random function can be constrained to produce the spatial variation of these three statistical parameters then that function can be used in equations (2) and (3) to simulate particle-fluid interactions for turbulent flows. Although the pseudo-random function will not represent any real flow, an attempt will be made to use it to estimate the accuracy of LV turbulent flow measurements.

Attempts to develop this pseudo-random function with simple periodic series summations have matched the mean velocity and the root mean square velocity, but do not match the auto-correlations function. Therefore, other approaches were investigated for developing the desired pseudo-random velocity representation.

Approaches to the Problem

In addition to the attempt to solve the problem with the simple periodic function, three other approaches were investigated this summer:

LSU Time Series Approach. Louisiana State University has developed a multiple input - single output time series model.[5] The model is structured as a linear combination of present and regressed values of the input series, a stochastic noise term, and a constant. Unfortunately, this method requires the desired pseudo-random function as an input, and cannot be used to derive this function.

Atmospheric Turbulence Model. Dr. Fichtl's paper deals with a model of the vertical non-homogeneous structure of turbulence in the atmospheric boundary layer.[6] This approach as developed by Fichtl and Perlmutter was also abandoned after discussions with Dr. Fichtl revealed that it could not be applied to this problem.

UTSI Turbulence Simulation. The final method investigated was developed by Frost and Perlmutter at the University of Tennessee Space Institute.[7] The approach is similar to the approach of Fichtl and Perlmutter, but simulates 3 components of pseudo-random turbulence from computer generated white noise. A number of white noise signals are passed through lagging and shaping filters so that the statistical parameters of the resulting signal have the predetermined first and second moments (mean, r.m.s and auto-correlation). The filters are determined following procedures outlined in references 6 and 7.

This method appears adequate for determining the desired pseudo-random turbulent velocity simulation. In addition, the method has the added advantage that it can be used to produce a time series which will give the particle motion in the simulated flow. This can be achieved by developing a second filter using equations (2) and (3) which will filter the simulated velocity signal.

References

1. Rudd, M. J., "A Self Aligning Laser Doppler Velocimeter," Optical Instruments and Techniques, 1969, pp 158-166.

2. Meek, C. C. & Jones, B. G., "Studies of the Behavior of Heavy Particles in Turbulent Fluid Flow," Journal of Atmospheric Sciences, Vol 30, No. 2, Mar 73, pp 239-244

3. Barnett, D. O., Private Communication.

4. Cybernetics, Norbert Wiener, John Wiley & Sons, Inc, New York.

5. Multiple Input - Single Output Time Series, John Tabony, Louisiana State University, Dept of Mech Engr, Dec 73.

6. Three Velocity Component Atmospheric Boundary Layer Turbulence Modeling, Frost and Perlmutter, Univ of Tenn Space Institute, 1976.

7. Random Data - Analysis and Measurement Procedures, Bendat & Piersol, John Wiley, New York.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78-0348	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 1977 USAF-ASEE SUMMER FACULTY RESEARCH PROGRAM VOLUME I		5. TYPE OF REPORT & PERIOD COVERED INTERIM 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J FRED O'BRIEN		8. CONTRACT OR GRANT NUMBER(s) F44620-75-C-0031
9. PERFORMING ORGANIZATION NAME AND ADDRESS AUBURN UNIVERSITY ENGINEERING EXTENSION SERVICE AUBURN, ALABAMA 30830		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2307D3 61102F
11. CONTROLLING OFFICE NAME AND ADDRESS AIR FORCE OFFICE OF SCIENTIFIC RESEARCH/NA BLDG 410 BOLLING AIR FORCE BASE, D C 20332		12. REPORT DATE Sep 77 13. NUMBER OF PAGES 1134
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

SUMMER FACULTY RESEARCH PROGRAM

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report represents the third years effort of an annual 10 week summer research program conducted by university faculty members at selected USAF System Command Laboratories. Program objectives are 1) To provide scientific and technological benefits to the USAF while enhancing the research interests and capabilities of engineering educators. 2) To stimulate continuing relations among participating faculty members and their professional peers at the AFSC laboratories. 3) To form the basis for continuing research of interest to the Air Force at the participant's institution. 4) To sponsor research in areas of mutual interest to the USAF, the faculty member, and his institution.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)